

When not to swim against the tide

Governments are habitually seeking to regulate the distribution of electromagnetic signals within and even outside their territory, but usually for reasons more concerned with wish-fulfilment than technological logic.

FIRST of all, a parable corrected: the King of England called Canute (995–1035) did not (repeat *not*) take up a position on the coast of southern England below the high-water mark so as to demonstrate to his people that he could keep the incoming tide at bay, but on the contrary to show them that even one as much revered as he could not influence natural phenomena. It would be a great help to us all if modern governments could follow Canute's modest example, especially in the exercise of their perennial ambition to mould technology to their wishes. So much should now be clear, in the United States and Western Europe, from the way in which tidy-minded arrangements for shaping the development of the telecommunications and television industries are everywhere collapsing because of developments that governments and their regulators failed to foresee.

Take, for example, the United States. Once upon a time, there was a telephone company called the Bell System (or, by those with fond feelings towards it, "Ma Bell") which had equipped itself to connect most telephone users in the United States to others like themselves and those living overseas as well. Because it had become a *de facto* monopoly, state and federal governments took powers to regulate the prices it charged for its services, but nobody was content with that. So the time came when the Bell System was broken up by the federal courts into several component parts: a handful of regional telephone companies (called "baby Bells") and an international carrier called AT&T, which naturally was constrained from competing by means of copper wires with the regional companies. What the courts could not have known was that cellular telephones based on radiotelephony would become feasible and fashionable, and that AT&T would buy (last month) a successful cellular telephone company called McCaw and be well placed to service a substantial fraction of the more profitable customers for telephone service in the United States.

A different kind of nonsense has been perpetrated in Britain, where the once-nationalized telephone monopoly has been turned into a private company owned by shareholders, many of them ordinary people. Instead of the US formula of disaggregation, the British government has chosen to regulate the *de facto* monopoly that stems from British Telecom's ownership of copper wire threaded to most British houses by direct regulation of its tariffs, by requiring that it should make its network accessible to competitors and by restricting its freedom, at least for the time being, to provide cable television to potential customers. Now, a full

decade after Mr Kenneth Baker, then a member of the British government, was promising "to cable" Britain in the succeeding ten minutes or so, the cable companies are cheerfully chipping away at British Telecom's telephone business.

In Western Europe as a whole, there is a greater muddle. Most telephone companies remain nationalized industries (and some, such as the German, are grossly inefficient). The distribution of encrypted television signals from satellites is growing quickly, as is cable television and cellular telephony. Governments have adopted a variety of regulations ranging from tight control to wishful thinking. Television broadcasts raise most hackles. Some governments regard foreign broadcasts as an assault on national culture, which conflicts with the European Commission's directive requiring the free flow of television signals within Europe. And that has the consequence that the British government, which has set its face against pornography on television, cannot regulate such signals spilling over from elsewhere if nobody requires that they should be paid for.

These are only some of the symptoms of the contradictions of principle still being built into national and international systems for regulating the distribution of electromagnetic signals of various kinds. Of course, it is entirely proper that commercial organizations should not be allowed to hold countries or international groupings to ransom by exploiting *de facto* monopolies established by the sheer volume of their present business, but what the sorry record of the past few decades has shown is that, at least in telecommunications, structural regulation is likely to be less effective than financial regulation. And even that may be less necessary than commonly supposed, given the apparently endless lode of technical innovation now being worked. That, of course, is the point that Canute was trying to make. □

Europe's nomenclature

The Maastricht Treaty has sown confusion in the names of European institutions.

THE celebration last weekend by Europe's political leaders of the entry into force of the Maastricht Treaty was properly modest. Winning ratification of the treaty has been a burden for most of them. It may even turn out to have been an albatross around some necks, a sign of electoral trouble



ahead perhaps. No wonder that the politicians chose not to mark the occasion with some great Eurospectacle, say Beethoven's Ninth Symphony on Ice, nor even with another confusing row about the Common Agricultural Policy. Yet confusion persists. What is the consortium of the twelve signatories of the Maastricht Treaty now to be called?

Confusion begins with the formal title of the treaty, which is the "Treaty on European Union". Its first sentence declares that its signatories establish among themselves a "European Union". So, henceforth, will the abbreviation "EU" stand for what used to be incorrectly called the "European Community" and correctly, if pedantically, the "European Communities", abbreviated as "EC"? One answer is "YES"; another, unfortunately, is "NO, not always". And the present confusion is inseparable from past confusions, and in particular from the coexistence of three separate European communities.

The first of these was the European Iron and Steel Community, the second that known as Euratom. Then in 1957 the Treaty of Rome created the third "community", called the European Economic Community (abbreviated as "EEC"). Neither Maastricht nor the Treaty of Rome abolishes its precursors, some of whose separate functions still persist. (The agreement between the International Atomic Energy Agency on the monitoring of fissile material in Western Europe is with an inspection agency established by the Euratom Treaty, for example.) But the Treaty of Rome also created the European Commission, which has become Europe's executive branch of government with important legislative functions; it promulgates European legislation approved by the European Council (representing member states' governments) and the European Parliament. Under Maastricht, the commission, the council and the parliament (not to mention the European Court of Justice and the Court of Auditors) have the same status with respect to each of the three precursor treaties.

So what is the difference between the "European Union" and the "European Community"? Formally, it must reside in the functions allowed under Maastricht, but not foreseen by the Treaty of Rome and its two precursors. (Maastricht's agreement on monetary union consists exclusively of amendments to the Treaty of Rome and so is the concern of the old EEC, now called the European Community.) These are the arrangements for evolving a common foreign and security policy and for collaboration on legal matters such as immigration and crime detection. On both issues, the European Council is in charge, the European Commission has a lesser role than usually. In other words, if there were to emerge a common policy on Bosnia (the chance is slim), that would be a business for the EU. So, too, would be a decision that policemen from one country could be drafted to another in emergencies. But a decision that coinage of the intended European currency should carry the head of, say, Charlemagne, would be one for the European Community (EC). So, too, would a decision that there should in future be a category of scientist known as "principal investigator" whose members would alone be entitled not to pay tax on grants received. It is hoped that so much will henceforth be crystal clear. □

Structural Biology

The appearance of another new journal linked with *Nature* offers great opportunities.

NEXT year, *Nature* will be launching a second sister journal, called *Nature Structural Biology*—a companion, as it were, for *Nature Genetics*. The purpose, an explanation of which follows, is primarily to provide an outlet for the many excellent contributions to structural biology which are sent to *Nature* for publication, but for which there is at present no room. But as in the relationship between *Nature* and *Nature Genetics*, it will be for authors and not editors to determine the eventual placing of their research articles.

Specifically, contributions in structural biology submitted for publication to *Nature* will be refereed in the usual way. As in the past, many will be published in the weekly journal. But authors of research articles judged technically sound but for which there is no room, or which appear more suited for a specialized audience, will be asked whether they wish their contribution to be considered for publication in *Nature Structural Biology*. There will be some advantages, notably the avoidance of further delay, but authors will have an absolute right to refuse. Readers of *Nature* proper will also be given an account each month of what *Nature Structural Biology* has to say for itself.

But why not leave that task to the many journals, some specialized and some less so, already active in the field? That is a natural question to which there are several answers. First, it may be of some benefit to the research community that the electronic techniques for rapid publication used by *Nature Genetics* should be applied in another field. Second, experience has shown that there is particular benefit in even specialized journals that mix together comment and research reports. Third, journals such as *Nature Structural Biology* are likely to be relevant, but in ways that cannot now be foreseen, to the electronic distribution (as distinct from publication) of research reports. But these sister journals bring to *Nature* and its readers the benefit that the distinction between what can and cannot be published is less arbitrary, even invidious. Especially in rapidly growing fields such as genetics and structural biology, it is disheartening that so much excellent research must be returned in the mail.

It is bound to get worse before it can get better. The causes of the present boom in structural biology are easily identified. Technical developments such as the crystallization of macromolecules, the automation of X-ray diffraction analysis, the provision of synchrotron light sources and even the scanning tunnelling microscope have helped to show how the function of many macromolecules is a function of their shape. The other side of that coin is that attempts to predict the real-life shapes of molecules of biological importance from a knowledge of the way they are put together from smaller units are no longer quixotic assaults on the impossible. *Nature Structural Biology* will, like *Nature* itself, take a special interest in the light that structural studies throw on the functioning of important molecules in biology. □

Science academies call for global goal of zero population growth

New Delhi. The scientists of Africa, the Republic of Ireland and the Vatican said no. Those from Japan and Argentina declared that the matter was "outside their competence". And academies from four other countries found reasons not to sign the statement on population that emerged from the first-ever summit of the world's scientific academies, held in New Delhi last week.

But the scientific academies of India and China — representing the world's two most heavily populated nations — were among 56 such institutions to endorse a 15-page statement calling for "zero population growth within the lifetime of our children", and

describing this as essential for dealing successfully with humanity's social, economic and environmental problems.

The so-called New Delhi statement was drafted by the Royal Society of London, the US National Academy of Sciences, the Royal Swedish Academy of Sciences and the Indian National Science Academy (INSA). It was sent in advance to all the world's academies, asking for their endorsement.

According to Prakash Tandon, a past president of INSA and spokesman for the meeting, the Irish and pontifical academies had rejected it outright, perhaps not surprisingly given the Pope's recent renewal of his criticisms of birth control.

The academies of Spain, Georgia, Armenia and Italy offered excuses for not signing, some saying their officials had been on vacation, others that they had not had time to read the document. The most explicit disagreement came from the African Academy of Sciences — one of the sponsors of the summit — which issued a dissenting note saying that overpopulation is not a problem in Africa (see box).

The joint statement urges governments to adopt an "integrated policy on population and sustainable development". It blames both rich and poor nations for adding to environmental damage, but acknowledges that development is a "legitimate expectation of less developed countries".

The statement claims that zero population growth is achievable, given time, political will and intelligent use of science and technology. But only if "appropriate policy decisions are taken now to bring about the requisite social change".

A key message of the summit was that fertility rates cannot be reduced merely by

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

India's population is likely to reach one billion by the year 2000.

providing more contraceptives. Family planning has to be seen as part of the broader "reproductive health service" that includes the elimination of sex-based inequalities, and the provision of basic requirements such as clean water and sanitation.

Some participants complained that the summit soft-pedalled on the relationship between population growth and use of resources. The final declaration calls for "a new ethic that eschews wasteful consumption". But Michael Chadwick, director of the Stockholm Environmental Institute, said he would have liked more quantification of the issues, and target-orientated goals for countries to commit themselves to.

Others criticized the document for not acknowledging the central role of women in all issues concerning population and development. But whatever its shortcomings, Zhao Qiguo, representing the Chinese academy, welcomed what he described as a "good document", adding that the question of resource use was "political and beyond the purview of scientists". **K. S. Jayaraman**

Why Africa said no

New Delhi. The African Academy of Sciences (AAS) refused to sign the final statement of last week's population summit of the world's scientific academies, which calls for a global reduction in fertility rates, saying that it does not consider overpopulation to be a problem in Africa.

"For Africa, population remains an important resource for development, without which the continent's natural resources will remain latent and unexploited", the academy, which was one of the 15 that sponsored the summit, said in a statement presented to the meeting.

According to J. K. Egunjobi, head of programmes at the AAS, who read out the note of dissent, the most important issue now facing Africa is development, not population growth.

"Human resource development must therefore form part of the population-resource issue in the forthcoming Cairo conference [on population control]", he told the summit.

Egunjobi says that, for most native Africans, marriage is important both for companionship and for procreation. "The [summit] statement completely ignores the fact that for certain parts of Africa, infertility is a major problem", he said.

The AAS statement says that the international economic environment in which Africa's development programmes are defined and executed is "an important variable in the population debate". It adds: "The contribution of the North to Africa's population predicament must be acknowledged in any suggestion as to how that situation is to be confronted." **K.S.J.**

NIH awards gene centre grants

San Francisco. Three US universities — the University of California, San Francisco (UCSF), the University of Pennsylvania and the University of Washington in Seattle — are to receive grants from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) to establish centres for developing the application of gene therapy to the treatment of cystic fibrosis.

The new grants, each of \$2.25 million a year over five years, are intended to bring together basic, pre-clinical and clinical research in the field, and have been specifically recommended by Congress as part of its efforts to promote the whole field of gene therapy research.

They come on top of a separate decision

by the National Heart, Lung and Blood Institute to award six universities a total of \$4 million each year for five years to support research into gene therapy for cystic fibrosis and other heart, lung and blood diseases.

Each centre being funded by the NIDDK is intended to act as a regional hub that the NIH hopes will strengthen gene therapy research across a variety of diseases and encourage innovative pilot studies.

In San Francisco, for example, more than 20 investigators drawn from UCSF and nearby UC, Davis, will develop transgenic animals, study the ways specialized cells function, analyse the molecular packaging of genes and improve techniques to transport genes into cells. **Sally Lehman**

Japan disappoints seekers of 'foresight'...

Tokyo. Japanese scientists and government officials have been puzzled by a steady flow of visitors from Britain in recent months asking about their experiences with "technology foresight" — a technique in which Japan is believed to excel, but of which most Japanese officials have never heard.

The UK government is engaged in a major exercise, outlined in the recent white paper (policy document) on science and technology, to make technology foresight a cornerstone of its efforts to focus research on areas where it can contribute most effectively to wealth creation.

With this in mind, the Prime Minister, John Major, the cabinet minister responsible for science, William Waldegrave, and last week the House of Commons Select Committee on Science and Technology, have each visited Tokyo to learn about Japan's expertise in the area.

But all of these British government officials have come away largely empty-handed because, contrary to popular myth, the Japanese government does not play a central role in guiding the planning of private industry.

The UK white paper cites the example of forecasts made every five years by Japan's Science and Technology Agency (STA). Since 1971, STA has carried out 'Delphi surveys' of researchers in industry, academic institutions and the government to predict developments in science and technology over the next 20–30 years.

The latest survey by STA's new National Institute of Science and Technology Policy (NISTEP) was completed at the end of last

year. It surveyed more than 2,400 people, and made predictions on a thousand or so topics. (In a Delphi survey, the results of a first questionnaire are returned to those who have answered the questions for a repeat of the exercise, on the basis that more accurate predictions can be obtained from an iterative process.)

Some of the predictions are startling in their optimism. For example, that there will be a cure for AIDS in 2006; that prediction will be possible a few days in advance of large earthquakes (magnitude 7 or more) in 2010; and that the mechanism of ageing will be elucidated in 2013.

The UK white paper claims that the STA/NISTEP surveys, despite taking a long-term perspective, have proved "reasonably accurate". But an assessment in 1991 by NISTEP of the first 1971 survey found that only 28 per cent of 530 predicted technologies were realized by 1991.

Many Japanese companies participate in the surveys and read the reports. But there is no evidence that they act on them. Hiroshi Okazaki, general manager of the engineering, planning and coordination division of NEC, says the value of the results of such surveys decreases the longer the range of the prediction.

For a range of three to five years, he says, results can usually be trusted. But beyond that, results become "very fuzzy". And for the 10–20 year period, results are not particularly useful at all.

Many Japanese electronics and automobile manufacturers have their own inter-

nal means of predicting developments a few years down the road. NEC has a 'core technology' programme where 30–50 leaders in the research and development and business sides of the company get together to predict future technologies and to discuss what is required to realize them. Thirty to forty core technologies are selected and resources are concentrated on their development.

Fujitsu regularly polls employees for ideas. For example, 15 years ago the company successfully predicted the importance of gallium arsenide devices with high electron mobility. The company now produces mass volumes of high electron mobility transistors.

But Fujitsu researchers also predicted the development of supercomputers based on Josephson junction devices. Despite much effort by Japanese electronics companies, they have yet to be realized.

The Japanese government, however, plays no role in these processes. Indeed many feel that the STA surveys can actually mislead government policy. The 1971 survey, for example, forecast that prediction of earthquakes of magnitude 6 or more would be possible at the prefectural government level by 1996. In the 1976 survey, the prediction became 2003, and in 1981 it shifted to 2006. In the 1986 survey, the goal posts were moved slightly, and those polled predicted that earthquakes of magnitude 7 or more would be predictable a few days in advance in 2006. And in the latest survey the prediction is 2010.

Throughout this period, STA and other government ministries, on the basis of this consensus among researchers, have poured billions of yen every year into an earthquake prediction programme. But the goal of earthquake prediction is no nearer.

In another example, the 1971 survey predicted that medium-sized nuclear-powered container ships would be in use in the mid-1980s. At that time, STA had just launched its experimental nuclear-powered ship *Mutsu*.

STA has sunk more than a billion dollars into development of *Mutsu*, which has only spent a few days at sea (see *Nature* 353, 594; 1991). Now the agency is spending several billion yen on converting the ship to diesel power so that it can be used as an oceanographic vessel. The latest STA survey predicts that nuclear-powered merchant ships will be in use in 2019. **David Swinbanks**

Correction:

The colour photograph used with our recent news item on Italian science was a portrait of Professor Enrico Garaci, president of the Italian National Research Council, and not as indicated that of Professor Luciano Maiani, director of the National Institute of Nuclear Physics. We apologize for the confusion.

... and UK Labour Party is unconvinced

London. Bipartisan support for the British government's plan to use a National Technology Foresight Programme as part of its new research strategy is unlikely to be forthcoming in the near future.

The question of whether reliance on this technique had been approved by the opposition parties in the House of Commons has been raised during the consultation process undertaken by the Office of Science and Technology (OST) in the past few weeks (see *Nature* 365, 776; 1993).

At a meeting organized by the OST at the Royal Society last week, for example, Sir Christopher Lewinton, chairman of the industrial conglomerate TI Group PLC, asked whether steps had been taken to ensure that a change of government would not precipitate a change of policy.

Professor Bill Stewart, the government's chief scientific adviser, replied that the government's plans were consistent with Labour Party statements on research. But he did not respond when questioned further on whether the party had, indeed,

been asked for its opinion.

In practice, Labour Party politicians seem divided. Some see no reason why they should not add their weight to the government's proposals. But Lewis Moonie (Labour, Kirkcaldy), the party's spokesman on science and technology, says that although the party supports the principle of trying to identify future technological trends, it feels that the government's approach is "sterile".

Moonie says that the Labour Party is keen to support technology foresight as part of an overall strategy for research. But he adds that there is nothing in the government's statements on research policy to show how the outcome of the foresight exercise will be used.

Moonie says he is concerned that the outcome of the exercise will be either forgotten or used as an excuse for shifting funds from basic research to areas defined as being strategically important. "What we have at present is a load of cobblers", he says — derived from Cockney rhyming slang for an impolite derogatory epithet. □

Protests greet new NSF rules on page charges

Washington. Professional societies have protested to the National Science Foundation (NSF) that a recent change in the foundation's rules, allowing research grants to help cover the costs of publication in commercial journals, will divert money into the pockets of private publishers.

The change has been under consideration since last winter, and took effect at the end of September. Scientists receiving NSF grants can now claim the cost of page charges for the publication of their work in all journals; previously they could do so only if the journals were not-for-profit, such as those published by learned societies.

Fred Spilhaus, executive director of the American Geophysical Union (AGU), described the change as "a perversion of the research funding process". The AGU, whose publications get about one-third of their income from page charges, says that the original system was introduced to help keep the subscription rates of learned journals within the reach of ordinary scientists.

The organization is angry that the change will enable money to be paid to high-subscription, commercial journals that circulate only to libraries. "It probably won't hurt us that much", says Spilhaus. "But my concern is a philosophical one: it's wrong for research money to be used in this way." The AGU has called on its 30,000 members to lobby the new NSF director, Neal Lane, to reverse the change.

Bob Hardy, contracts division director at the foundation, says that the decision brings NSF into line with other agencies, such as the National Institutes of Health. "Our basic consideration was that the principal investigator should be allowed to make the decision", he says.

An estimated 0.6 per cent of NSF grants, or about \$20 million, is allocated each year for the publication of research findings. The decision to include commercial publications was taken in July, before Lane's confirmation. But Hardy says the change in policy was confirmed by the new director.

Not all learned societies are worried about the change. A spokesman for the American Physical Society said that the APS welcomed the "level playing field", and doubted if any commercial publishers would introduce page charges to take advantage of it.

But Gordon and Breach, a commercial publisher with more than a hundred journals in the physical sciences, has announced a scheme under which investigators who pay a voluntary page charge on publication of their work will receive, in return, vouchers of equivalent value for themselves or their library to spend on Gordon and Breach publications.

Colin Macilwain

Opponents of 'earmarking' claim victory in Congress

Washington. Critics of the way in which the US Congress allocates funds to science projects favoured by members of powerful congressional committees are claiming credit for the fact that the amount of money earmarked in such a way has been cut drastically in this year's budget.

Years of sharp growth in 'pork-barrel' funding saw the total money allocated to science projects reach \$760 million last year, most of it involving the construction of new facilities at universities (see *Nature* 364, 563; 1993). But an early assessment of the budget bills for the current fiscal year, which started on 1 October, suggests that the practice has been sharply curtailed.

George Brown (Democrat, California), who has led a high-profile campaign against earmarking in his role as chairman of the House of Representatives Science and Technology Committee, says that the amount allocated by two key budget committees to such projects has been cut by four-fifths.



Brown: campaign continues

In principle, the main goal of the campaign against earmarking has been to improve the quality of publicly funded research by ensuring that all projects supported by the government are properly peer reviewed. In practice, however, the issue has also become part of a wider struggle for power in the House of Representatives between younger members who sit on committees such as Brown's, and the more powerful (if less numerous) members who occupy positions of influence on the appropriations committees, which determine agency budgets.

The balance of power appears to be shifting, at least insofar as the science budget is concerned. Earmarked funds in the energy and water section of the budget, which includes research projects financed through

the Department of Energy, have fallen from \$180 million to \$40 million. Those in the part of the budget bill covering independent agencies, such as NASA and the National Science Foundation (NSF), fell from \$60 million to just \$8 million.

"People in the science community who were looking for a downturn in this [practice] are going to be happy", says Brown. But he admits that the earmarking process is unlikely to disappear entirely, and pledges to continue his fight against it.

The assessment by Brown's staff defined earmarked projects as budget expenditures whose location had been specified in the budget bills, but had not been specifically authorized by Congress, nor requested by the president. Their analysis did not include defence spending, traditionally a big source of earmarked funds, but for which a budget has not yet been finalized.

The staff assessment showed that earmarking in another part of the budget — that covering the Departments of Commerce, Justice and State — has crept up from \$76 million to \$83 million. Overall, however, the amount of earmarked science in the four sections studied had been halved from \$410 million to just under \$200 million.

Brown's staff stress that their analysis of this year's budget is provisional, and admit that earmarked funds can exist deeply hidden inside the complex budget bills. But they claim that even these provisional results are sufficient to indicate a degree of success for his campaign.

Brown's supporters must now ensure that agencies such as the NSF receive enough money to build properly peer-reviewed research facilities. Recipients of 'pork' often point out that it is the only source of federal funding for new facilities, since agency budgets for this remain tightly restricted.

This year's doubling of the NSF's facilities budget, to \$100 million, will help to correct this situation (see *Nature* 365, 467; 1993). But more money will be required next year to compensate for that lost by universities in earmarked funds.

Colin Macilwain

Schwitters steps down as SSC director

Washington. Roy Schwitters has resigned as director of the Superconducting Super Collider (SSC) Laboratory, which has been directing the construction of the huge particle physics project cancelled by Congress two weeks ago.

Schwitters has borne the brunt of public criticism of the management of the project. "I'm interested in the science, not in manag-

ing the demobilization of the laboratory", he says. "I want to get on with thinking about the future." The former Harvard physics professor says he will continue in his post in order to redirect the science there, until the Universities Research Association, which manages the laboratory for the Department of Energy, appoints a successor.

Colin Macilwain

'Publication by press conference' under fire

Paris. Worldwide media coverage of claims by researchers at the Pasteur Institute in Paris that they have identified the 'door' through which HIV infects cells — thus opening the way to possible new treatments and a vaccine for AIDS — has led to renewed calls from scientists for their colleagues to stop announcing 'break-throughs' at press conferences before they have published their data.

Ara Hovanessian and his colleagues at the institute claimed at a press conference in Paris that CD26, an enzyme that is specific for dipeptide bonds containing the amino acid proline, and which is found on the surface of human T cells (the targets for HIV infection), is essential for assisting the entry of HIV together with the T-cell surface receptor CD4.

According to the French group, CD26 cleaves the HIV envelope protein gp120 at the V3 loop region as the virus attaches to the cell surface. They claim that enzyme inhibitors and antibodies to CD26 block the entry of HIV-1 and HIV-2 into human cells, and that expression of both CD4 and CD2 is needed for virus entry into mouse cells that do not normally express these molecules.

The claims, which are made in a paper that has been submitted to *Science*, were made public at a press conference on the first

day of the three-day "Cent Gardes" AIDS colloquium at Marne-la-Coquette near Paris. They immediately captured headlines worldwide (including major stories in *Le Monde* and *The New York Times* and the London *Evening Standard*).

Scientists find a new key to Aids

FRENCH scientists who discovered how the Aids virus penetrates body cells have made a major breakthrough, opening the way to a possible vaccine.

Luc Montagnier, discoverer of the deadly virus, announced today: "This is a

STANDARD FOREIGN
NEWS DESK

department of virology and cellular immunology, said his research team had isolated a co-receptor, labelled

we will have vaccines against different types of the infection."

A team has already begun experimenting on mice.

The institute said the discovery of CD26's major role in the mechanism allowing the virus to penetrate and

From the London *Evening Standard* of 26 October.

Hovanessian says that it was "unfortunate" that his findings were given such wide exposure and maintains that he had no choice but to present his findings at the colloquium as they had already been "leaked".

But Hovanessian also agrees that he gave an interview to the news agency Agence France Presse several hours ahead of the press conference. Both Hovanessian and Luc Montagnier, the head of the Pasteur's AIDS research unit, were quoted in newspaper articles as describing the finding as a major breakthrough that could lead to new

treatments and vaccines. Both also defend the decision to go public, saying that information on AIDS should be released quickly.

But Simon Wain-Hobson, a researcher working in another AIDS group at the Pasteur Institute, describes the data that Hovanessian presented on the second day of the conference as "totally unconvincing". Paul Maddon, chief executive officer of the New York company Progenics Pharmaceuticals Inc. — and who cloned the gene for CD4 and showed that it was the receptor for HIV — says that "it suggests a role for CD26 but is not definitive proof".

Both Wain-Hobson and Maddon are concerned that, if CD26 is shown to have no role in HIV infection, it will harm both the public and the credibility of AIDS research. Such credibility, they say, is already low because of the anticlimax that followed widespread publicity of "rumours" of the success of the Salk therapeutic vaccine marketed by Immune Response Corporation (as well as the publication in *Nature* of a multidrug therapeutic strategy which was later retracted).

Hovanessian and Montagnier are both staking their reputations on CD26. Hovanessian admits that he failed to convince his audience at the conference: "We were convinced ourselves. It wasn't a matter of convincing [others]; it was a matter of believing." Montagnier says he has read carefully the manuscript submitted to *Science*. "I am very convinced", he says.

But Wain-Hobson argues that the judgement of scientific results should not be left to individual scientists, particularly in AIDS research. His prediction that similar incidents would occur within twelve months was borne out even sooner. Last weekend, a group in Marseilles made headlines after announcing at a press conference that they had discovered a peptide that blocked entry of HIV into cells *in vitro*. **Declan Butler**

MRC to limit patents on cDNA sequences

London. Britain's Medical Research Council has announced that it will no longer seek to patent cDNA sequences of unknown function. It claims that its initial bid for patent protection on these sequences was made simply to protect British interests if an earlier application from the US National Institutes of Health (NIH) had proved to be successful.

The original bid from the NIH provoked widespread protests from those in the scientific community who argued that it would immediately turn all knowledge of the human genome into private property. In summer last year, however, the latter application was rejected by the US Patent Office on the grounds that it did not meet the necessary criteria for patentability.

The MRC says that it intends to carry on with its own original patent application. In particular, it is keen to receive a ruling from the European Patent Office on whether the application is legally acceptable. And it is also continuing to encourage patent applications on gene sequences where the function is known (a move which itself is still criticized by some scientists).

But in order to show support for the US Patent Office decision, the MRC says that no new patent application will be filed "at

present" for sequences where the function is not known. As a result, all sequences obtained at its Human Genome Resource Centre will be transferred directly into DNA databases.

David Owen, director of technology transfer for the MRC, says that he hopes rulings from the US and European patent offices will discourage further applications based only on sequence data. **D.D.**

German stalemate on nuclear power

Munich. The future of nuclear energy in Germany has been threatened by the breakdown last week of negotiations between the coalition government and the opposition Social Democrats.

The national executive of the Social Democrat Party (SPD) rejected a compromise worked out by the environment minister, Klaus Töpfer, and the SPD's Gerhard Schröder that would have allowed the development of extra-safe nuclear reactors, even though the SPD had agreed in 1986 to a complete non-nuclear energy policy.

The SDP has now accused Schröder of trying to sell his own party an "ambivalent

strategy". The coalition government in turn accuses the SPD of undermining a fundamental aspect of the German economy; the country's ailing coal industry depends on subsidies from the nuclear industry and from nuclear energy generators such as RWE, Veba and Siemens KMU.

Negotiations to establish a working life of between 35 and 40 years for existing reactors are also being blocked by the SPD, which objects to the so-called 'long-term interim storage' of nuclear waste. The stalemate means that a new law is unlikely to be passed until after the general elections next October. **Matthew Beard**

Maastricht lets parliament flex its muscles

Brussels. The European Parliament wants to use the new powers given to it when the Maastricht agreement came into force on 1 November to play a bigger role in shaping Community research policy. Previously, control has been exercised mainly by the European Commission, which proposes research programmes, and the Council of Ministers which approves them.

Proposals have been required to go through a double reading in the parliament. But the commission has been able to ignore any amendments the parliament asked for; the only way the parliament has been able to change the programmes has been to threaten to use the double-reading procedure to delay their adoption.

The main new power given to the parliament is that, like individual member states, it can now veto the Framework budget proposal (see *Nature* 365, 775; 1993). But members of the parliament are keen to dispel the idea that this restricts them to wielding a very large club. Their power of veto, they argue, carries with it the right to shape EC policy from the outset.

If parliament is to do this, however, it must first break the grip that the commission holds over Community research as the body that both proposes and manages programmes. Officials at parliament admit that it does not have the expertise to take over the commission's role in proposing programmes. But it wants, for example, to stop the self-perpetuating tendency of unproductive programmes.

In particular, the parliament is insisting that the commission takes into account evaluations of the fourth five-year Framework programme, which is meant to start next year made by "external, independent experts qualified in research management", before proposing its successor.

The parliament also wants a bigger say in overseeing research programmes. At present, decisions on who gets what are made in secret by committees composed of national civil servants and chaired by commission officials — a procedure that actively encourages nationalist horsetrading, says one parliamentary official.

Members of parliament were last week informally discussing whether their "powers of scrutiny and control" should extend to their direct participation in these committees. They also say that the committees should meet in public when they are not discussing confidential matters, but only

what sort of research to do.

Among the changes that the parliament wants is an increase in research into the social and economic impact of research in the fourth Framework, and an ambitious programme on these topics in the fifth.

Parliament is also broadly in favour of spending more money on life sciences, environment, renewable energy and new industrial technologies, and less on the costly information and communications technology programmes.

Claude Desama, the president of parliament's energy, research and technology committee, says that its relations with the commission have improved since Antonio Ruberti replaced Filippo Pandolfi as research commissioner at the beginning of this year (see *Nature* 361, 198; 1993).

In particular, he says that the parliament welcomes Ruberti's greater emphasis on fundamental research, the closer integration

of Community programmes with those of national and European research bodies, and its growing attention to the social and economic dimensions of research which it would like increased.

But the parliament feels that Ruberti needs to do more to make the commission's directorates of research and industry cooperate. It also believes that Ruberti has missed an opportunity in the fourth Framework programme to reform the Joint Research Centres, despite long-standing criticism of their relatively low productivity.

If the parliament is to play a greater role in EC research, it will first have to address its lack of relevant infrastructure and expertise, highlighted in a report last year by Michel Hervé, a French member of the parliament. He also recommended that the parliament be given access to the same information as is currently available to the commission.

Declan Butler

CERN members seek better deal

Munich. The European Laboratory for Particle Physics (CERN) is considering a proposal that could result in the level of industrial contracts received by its 17 member states being more closely related to their financial contributions.

Last month, Italy formally suggested how such contracts, most of which go to the two host countries, France and Switzerland, could be more equally shared without increasing the cost to CERN. This would be done by giving non-host countries an opportunity to match host-country bids. A working group within CERN's financial committee will make recommendations to the next general council meeting on 17 December.

Contributions to CERN are related to each country's gross national product. The number of scientific staff from each member state is unrelated to the contribution of that country, and construction and servicing contracts are awarded to the lowest bidder.

Some countries, however, feel that the way this works out in practice has not been equitable. In particular, they claim that CERN's location on the border between France and Switzerland gives them a strong advantage when bidding for contracts.

Between 1986 and 1991, for example, France, which currently provides 17 per cent of the laboratory's annual budget of SFr954 million (US\$640 million), received 33 per cent of all contracts. Similarly Switzerland, which contributes less than 4 per cent, received 24 per cent of contracts. In contrast, Germany (which pays 23 per cent of the budget) received only 13 per cent of contracts, the UK (14 per cent) received 7 per cent, Italy (17 per cent) received 6 per

cent, and Spain (8 per cent) only 2 per cent.

Italy's delegate on the CERN council, Luciano Maiani, director of Italy's National Institute of Nuclear Physics, says that contracts should still be awarded on the basis of cost but that the system of assessing tenders should give non-member states a fairer chance of winning.

At present, technical specifications are sent to member states and, when sealed tenders are opened, the contract is given to the lowest bidder. Maiani suggests that non-host countries be given a chance to match a lower bid from a host country; if they can meet it, they should get the contract.

Spain has been the most active member in protesting at the disparity between contributions and returns. It has not paid its membership dues for nearly two years, insisting that either these be reduced, or it should be guaranteed more contracts, while stressing that it is keen to remain an active member.

Spanish scientists have been embarrassed by their government's action, and have been pressing the ministry of industry to pay the country's debts. Later this month Carlo Rubbia, CERN's director general, and William Mitchell, the council's president, will visit Madrid to negotiate repayment.

Cayetano Lopez, Spain's representative on CERN's council, accepts that Spain has a responsibility to pay off its debt in a reasonable time. But he also argues that it is necessary to work out a fairer deal.

Spain is also unhappy that it has less than two per cent of CERN's 3,000 permanent staff. But the principle of hiring staff according to merit rather than country of origin is not under discussion. **Alison Abbott**



Ruberti: approval from parliament.

Spanish laboratory backs off bid to cut research staff

Madrid. Spain's Council for Scientific Research (CSIC) last week backed down on a decision to eliminate the posts of 14 researchers from the Biological Research Centre (CIB) in Madrid, the largest of its hundred or so institutes.

The council has apparently bowed to pressure from the scientists to reverse a decision taken after two independent evaluations had given them low scores for productivity, and the former director of the institute had recommended that they be "relocated" elsewhere.

The decision is a blow to CSIC's attempts to boost the productivity of its centres by penalizing scientists whose output is felt to be unacceptably low. Last February, the council changed its rules to increase the power of institute directors over the way their laboratories are organized, and to reduce the involvement of staff in management decisions.

The impact of these changes has already been felt by researchers at the Cajal Institute of Neurobiology in Madrid, whose 160 staff members include 35 scientists. After an evaluation in January 1992, the institute was reorganized into four departments, and four researchers whose productivity was considered to be very low lost some of their technical staff and laboratory space.

The centre's director, Alberto Ferrús, says that the decision was painful, but was necessary to improve the scientific performance of the whole institute. Similar changes are now feared at the CIB, which has 500 staff, including 86 scientists. The centre was widely known in the 1960s and 1970s for its work in enzymology and molecular biology, but since then its scientific productivity has not grown as fast as hoped.

In March 1992, an external evaluation committee complained that the institute had no strategic plan, was split into groups that were too small to carry out internationally competitive work, and did not have enough young scientists.

In response, the institute's then acting director, Manolo Espinosa, announced plans to reorganize CIB into five departments. One proposal, apparently backed by CSIC, was that the positions of 14 researchers would disappear because, it was claimed, their work did not fit into the new scheme.

Most of these scientists had been in charge of small groups whose performance was considered by the evaluation committees to be particularly low. Between them they brought in less than 5 per cent of the institute's grant income and published far less than other scientists on the centre's staff. Espinosa proposed that they should be transferred to other CSIC institutes.

But the scientists involved challenged the report's conclusion, backed by the Association of CSIC Staff Researchers, the professional association to which 30 per cent of CSIC's researchers are affiliated.

The controversy provoked a second evaluation by a new committee in July. This supported the original committee's conclusions. But the CIB scientists continued to oppose Espinosa's plans for reorganization, and led a public campaign against them.

The situation changed in the last week of September when Espinosa's temporary contract expired and a new director, Guillermo Jimenez, who had worked for several years in the institute, was appointed.

Jimenez revoked Espinosa's plans — once again with the backing of CSIC — and redistributed the 14 researchers within the new departments, allowing them to keep the laboratory space and technical assistance they had enjoyed before the evaluations.

The decisions of the new director are widely seen as CSIC's response to pressure from the scientists' union. But the changes recently introduced by CSIC in the way that research institutes are run in Spain are still considered by many scientists necessary to increase their effectiveness.

The new rules, introduced earlier this year after seven years of debate, increase the power of the directors, as well as of the president of the CSIC itself and its advisory boards. The rules also reduce the presence of researchers on CSIC's 24-member scientific and executive advisory boards.

The latter move has itself provoked widespread criticism from some researchers, who claim that the institutes are becoming less democratic. But others feel that what has happened at CIB shows that the reorganization process is itself in danger of becoming ineffective.

Oscar G. Segurado

Snapshot of French science shows growing impact

Paris. Spending on science in France increased from 1.97 to 2.4 per cent of the country's gross national product during the 1980s, a reflection of the priority given to research by the Socialist government which came to power in 1981, according to the second report of the Observatoire des Sciences et des Techniques (OST).

This increase in funding was accompanied by a 10 per cent rise in France's share of publications in international journals to 4.7 per cent (although some of this increase may be accounted for by the trend to publish more in English). In contrast, over the same period, the United Kingdom's share fell by 8 per cent to 8.4 per cent, and Germany's by 5 per cent to 6.2 per cent.

The OST report is the French equivalent of the Science and Engineering Indicators produced by the National Science Board in the United States, with more than 350 tables of statistics on French research and its position compared to the rest of the world.

If the report provides evidence of the impact of increased spending on research during the 1980s, it also provides information that is likely to help politicians to decide where best to reduce spending in the country's current economic crisis.

In particular, because the breakdown of public spending on science is by theme rather than by research organizations, it is able to reveal the continued preoccupation with defence and strategic technologies which began with de Gaulle in the 1950s.

More than half the FF86 billion (US\$14.6 billion) public spending on both civilian and military research goes on four large programmes: space, aeronautics, telecommunications and nuclear. The clear implication is that the government might do better to reduce its spending on large programmes rather than imposing, as it has done, almost zero-budget growth on research organizations in next year's budget (see *Nature* 365, 598; 1993).

Rémi Barre, the director of OST, says that the high level of spending on these large programmes also challenges the widely held assumption that France spends more than other European countries of its science budget on public sector projects.

Health, transport and the environment, for example, accounted for less than 10 per cent of the total national spending on research by both government and industry (FF158.7 billion). Similarly, funding on technology transfer and innovation over the same period amounted to just one per cent of total spending on science. **Declan Butler**

Science & Technologie Indicateurs 1994, (OST, 93 rue du Vaugirard, 75014, Paris)



Madrid's Biological Research Centre

Battle looms over future of Science Museum

London. A sharp conflict has broken out over the future plans for Britain's Science Museum. In one corner are traditional historians of science and technology, who want to see the museum preserved primarily as a national store and display case of historical artefacts. In the other are those who argue that more of the funds devoted to these activities should be diverted towards encouraging greater public understanding of contemporary science.

The trigger for the conflict has been the publication last month by the museum's director, Neil Cossons, of a new development plan. These include an increase in capital investment — for example in new galleries and other facilities designed to encourage greater public access to the museum — but also a reduction of 25 per cent in the number of curators.

Cossons defends this shift in the museum's priorities as reflecting the changing environment in which the museum operates. In particular, he says, it is in line with the government's desire to see greater effort being put into the dissemination of information about modern scientific and technological achievements.

The Science Museum has, for example, already been identified in the recent white paper on science as the site of a major government-sponsored exhibition on contemporary science to be held in the year 2001, celebrating both the beginning of the new millennium and the Great Exhibition of 1851.

But such moves are being viewed with

the History of Science.

Last week, the society issued a statement criticizing, for example, the removal of a historical display of electricity and magnetism — including apparatus belonging to the nineteenth century physicist Michael Faraday — to make room for a celebration of modern medicine.

But John Durant, who holds a joint appointment as assistant director of the museum and professor of the public understanding of science at nearby Imperial College, and is widely seen as one of the architects of the proposed changes, justifies the shift in direction as matching the government's published ambitions in encouraging the greater public understanding of contemporary science.

"The challenge is to rethink the museum's historic mission in ways that are relevant to the rapidly changing circumstances of science and technology in the 1990s", said Durant, writing in a recent issue of *Museums Journal*.

Voluntary redundancies have already

been requested from existing staff, and Cossons says he hopes that much of the reduction he is looking for will be achieved in this way. But he does not rule out the possibility that some staff will have to leave.

Any such involuntary cuts are being strongly resisted by the various unions representing the staff of the museum, in particular the Institute of Professional Managers and Specialists, to which most of the curators belong.

The unions are already balloting their members on strike action. And in a joint letter to the *Museums Journal* the staff representatives of five different unions at the museum dispute that the cuts are needed to meet the changes outlined by Durant in his article. "At a time when the importance of producing effective displays which enhance the public's understanding of science and technology has never been greater, the Science Museum intends to make redundant the very staff who have the knowledge and experience to accomplish this difficult task", say the unions.

David Dickson

Celltech announces public offering

London. Celltech, Britain's oldest biopharmaceutical start-up company, has announced plans to raise more than \$30 million by offering shares to the public through the London Stock Exchange. The new funds are to be used to finance existing and future development programmes and research projects.

Recent changes to the rules governing Stock Exchange listings for biotechnology companies in the United Kingdom have created the opportunity for a number of companies to meet their needs for funds by selling shares to the public.

"A Stock Exchange listing will enable us to raise finance to fund our existing and future projects," says Celltech's chief executive, Peter Fellner. "It will be our first such fund-raising since 1987 and will help fulfil our objective of developing into a profitable pharmaceutical company towards the end of this decade."

Indeed, Celltech has been able to fund its progress in recent years through deals with established pharmaceutical companies such as American Cyanamid, Bayer and Schering-Plough.

Most recently, Celltech signed a manufacturing and investment agreement with Hoffmann-La Roche which is expected to result in sales of up to £30 million in the next four years, and involves the Swiss company investing £5 million in Celltech.

Industry sources suggest that the Celltech Group is unlikely to make profits until the middle of the second half of the decade, although one of its two main businesses is

already in the black. Celltech Biologics, which provides contract manufacturing services to the pharmaceutical industry, including monoclonal antibodies, has been consistently profitable, and sales in 1992 were £12.9 million.

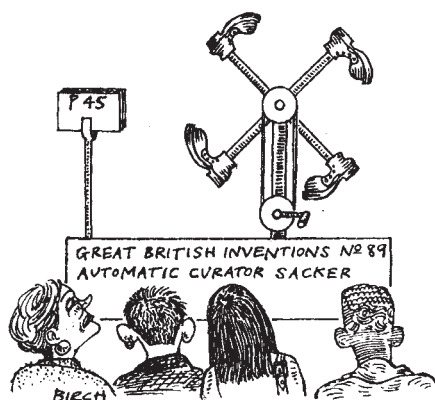
While Celltech's avowed intention is to become a fully integrated pharmaceutical company, many of its frontline products are being developed in close collaboration with other pharmaceutical companies. The most advanced product involving Celltech is a mouse-derived anti-tumour necrosis factor antibody belonging to Bayer which is in phase III clinical trials for septic shock.

Bayer is also interested in Celltech's novel anti-tumour necrosis factor antibody, CDP571, which is in phase II trials to treat septic shock and rheumatoid arthritis and also a phase I trial for inflammatory bowel disease.

American Cyanamid is linked with Celltech's next most advanced product, a calicheamicin-antibody conjugate, which is in phase I trials against ovarian cancer. Similar calicheamicin-antibody conjugates are in preclinical trials against acute myeloid leukaemia and colorectal cancer and also involve American Cyanamid, the discoverer of the toxic cell killer which is guided to the cancer cells by the antibody.

One of the main reasons for the public flotation, seen by many biotechnology companies as a rite of passage, is that Celltech is keen that its earliest investors (the company was founded in 1980) should see some return on their investment.

Mike Ward



concern by many historians of science and technology, who fear that job reductions among curators will mean less effort being placed on the preservation and interpretation of historical artefacts, and that the museum's role as a custodian of these is in danger of becoming sacrificed to the short-term priorities of government and industry.

"There is a balance to be struck between these two objectives, and some of us fear that the pendulum is swinging too far", says Geoffrey Cantor, professor of the history of science at the University of Leeds, and president of the British Society for

Irish science needs a champion

SIR — Your articles on the funding of basic scientific research in Ireland¹ highlight the bankruptcy of government policy on science, the funding of scientific research, the education of scientists in Ireland to postgraduate level and, most importantly, they emphasize the alarming degree of anti-scientific culture among contemporary politicians in this country.

As your articles correctly pointed out, Irish science badly needs a champion in government capable of taking a long-term view of the role of science and how it might flourish through orderly planning and a guaranteed injection of even modest funding. For too long the scientific community has had to survive from year to year on the basis of whimsical policy or none at all. Ministers in successive governments with responsibility for science and technology have largely chosen to ignore the scientific aspects of the brief in favour of the technological, often deliberately downgrading the status of science in the process (a notable exception was Michael Smith). Irish politicians and senior civil servants with responsibility for industry, technology and science might broaden their perspectives on the essential role of basic research and its value to a nation through reading Leon Lederman's thought-provoking articles^{2,3}.

I would like however, to correct the implication in "Irish funding disappears", which seems to suggest that EOLAS, the Irish Scientific Funding Agency, is somehow responsible in part for the situation in which Irish science finds itself. This is not so. EOLAS has served science very well, given its limited resources and the breadth and depth of its responsibilities. It has been a long-time champion of postgraduate funding for research and has successfully defended and administered various programmes of funding for research over the past decade. Its chairman, Professor Gerry Wrixon, has many times gone on record to state that "any sustained effort in applied research is useless unless you continue to fund basic research at the same time". It seems incongruous that EOLAS is about to be dismantled and subsumed into a new agency, FORFAS, where its commitment to scientific endeavour will be even more diluted. Yet another example of the devaluation of science and scientists in this country.

Ireland has a long and successful track record in attracting new industries, many with state-of-the-art technology. However, times are now changing because of strong competition from other markets. Confidence of potential investors must surely be shaken by the prospect of setting up high-technology industry in a country where less government money is invested in basic research than in sport and where

postgraduate remuneration is worth less than the minimum unemployment benefit. The irony is that funding of basic research in Ireland would not be very costly. An annual allocation of £3 million for funding in natural sciences, awarded on the basis of peer review, would broadly satisfy the needs of the research community. An investment of half as much again would support 500 postgraduate students to the tune of £3,000 each per annum, socially a much more acceptable scenario than the dole queue.

D. J. Fegan

Department of Experimental Physics,
University College Dublin,
Dublin 4, Ireland

1. *Nature* **364**, 658 & 662 (1993).
2. Lederman, L. *Scientific American* **252** (5), 34 (1984).
3. Lederman, L. *Science* **237**, 1125 (1987).

Words, idle words

SIR — I dissent from the argument by Ernest H. Williams Jr *et al.* (*Nature* **364**, 664; 1993) that the use of the word 'epidemic' should be confined to human diseases. A reference to the etymology of the term "Epidemiology" would have clarified the issue raised.

The word *demos* (Greek: *δημος*) has several meanings: notion of place — land inhabited by people, properly the part of the territory belonging to a community, thus region, country, people; notion of persons — (i) in the ethnic sense: population of a country, people, (ii) in the political sense: people, citizens, (iii) in a wider sense: people, species, breed (in general) in referring to animals, the species of birds, monkey etc.; both political and local notion — in Athens administrative division or district, subdivision of the tribe.

The latter meaning corresponds exactly to the Sanskrit root *dami* (the source of the Greek word *demos*), which signifies administrative division. It is, therefore, clear that the word *demos* contains notions of place, human population and animal population. Further, the general methodology in the field of epidemiology — the fundamental principles, the importance of statistics for the study of the populations and so on — is identical irrespective of whether the population examined is of animals, humans, plants or microbes. Thus, both from an etymological and a methodological point of view, there are no valid reasons for arriving at a restrictive meaning of the word "epidemiology" and monopolizing its use for human pathology.

According to *Dorland's Medical Dictionary*², a symptom is "any subjective evidence of disease of a patient's condi-

tion, i.e. such evidence as perceived by the patient; a change in patient's condition indicative of some bodily or mental state." The veterinary surgeon's assessments of a patient's health are frequently based on clinical evidence which fits the definition of symptoms cited above.

It is therefore not appropriate to be anthropocentric and to make distinctions between humans and nonhumans when using these terms in scientific publications.

S. K. Sharma

CSA University of Agriculture
and Technology,
Mathura Campus — 281 001, India

1. *OIE Bulletin* 95(5), 51–52 (1983).
2. *Dorland's Illustrated Medical Dictionary* 26th edn (Saunders, Philadelphia, 1981).

Free trade error

SIR — Your leading article in support of the North American Free Trade Agreement between the United States and Mexico (*Nature* **365**, 8; 1993) errs on one point. The classical doctrine of free trade assumes that capital isn't mobile across national boundaries. Or, as David Ricardo, the nineteenth century British economist who originated the doctrine, put it: "The difference . . . between a single country and many is easily accounted for, by considering the difficulty with which capital moves from one country to another. . ."¹

Remove that assumption and one can no longer argue that the majority of people in a rich country such as the United States, where the ratio of capital invested per worker is high, will automatically be made better off when trade barriers are relaxed with a relatively poor country such as Mexico, where the ratio of capital per worker is low. The reason is that trade barriers can act as an obstacle to capital mobility, especially in cases in which the poor country is intended to be used as an "export platform" by investors from the rich country.

That even economists are sometimes confused on this issue was pointed out by one of America's leading trade theorists, Columbia University's Jagdish Bhagwati: "[W]hereas free trade is mutually beneficial to all freely participating economic agents, this cannot be asserted for free international mobility of capital when free trade obtains initially. In popular discussions and in scholarly discourse the two propositions are put on the same level of acceptability, but that is simply an error."²

Luke Lea

818 Fairmount Avenue,
Signal Mountain, Tennessee 37377, USA

1. *The Works and Correspondence of David Ricardo*, Vol. I, On the Principles of Political Economy and Taxation 135 (Cambridge University Press, 1966).
2. *Lectures on International Trade* 298 (MIT Press, 1983).

Critics condemn NIH women's study

Nobody disputes that women have health problems that differ from those of men, but NIH may be wasting \$625 million or more in pursuit of an ambition to be seen to be politically correct.

WHEN the US National Institutes of Health (NIH) embarked two years ago on a \$625-million 14-year study of women's health, they were able for a time to bask in approval (or self-approval) for having done the "right thing". But now the Women's Health Initiative (WHI) has been severely criticized by a committee of the Institute of Medicine (IOM), itself appointed in response to a request by the Congress for an appraisal of one of NIH's most expensive programmes ever. The question now is what went wrong.

By any standards, WHI is a massive epidemiological trial. It is one of the largest and longest ever undertaken by the NIH. Eventually, some 160,000 post-menopausal women will have been asked to take part. Among the objectives are to tell whether low-fat diets will prevent breast cancer, whether hormone replacement therapy will reduce the incidence of coronary heart disease among women and whether calcium and/or vitamin D supplements will reduce the incidence of osteoporosis. The study was put together with unusual speed in response to the determination of Bernardine Healy, the former director of NIH, to join the women's health bandwagon.

The committee's appraisal of the study, published earlier this week, is a scathing indictment of its conception and design. The head of the committee, Marion J. Finkel of the Sandoz Pharmaceutical Corporation, was chosen specifically because she is in no way reliant on NIH grants. Among public documents of this kind, the committee's report stands out for its candour. The preface, for example, says: "Had this committee been asked to design a plan for women's health research, it would not have designed this Women's Health Initiative."

The committee criticizes the WHI on several grounds, including the following: (1) The study is based on premises that are scientifically weak; (2) it is unlikely to provide statistically valid answers to the questions it asks, partly because its statistical design is poor and partly because the questions it asks are scientifically limited; (3) the study's informed consent procedures do not adequately inform women of what they are getting into; (4) the \$625-million price, high as it is, probably underestimates the true cost.

The report makes clear the misgivings of the committee about the WHI. Thus the committee attributes the request for an evaluation of WHI to "nagging doubts ... that political concerns influenced the design, timetable and budget" on the part of the

House of Representatives Appropriations Committee. Bluntly, the House asked whether the study is worth the money it will cost. In its preface, the committee frankly admits that some members would have been happy to see the WHI cancelled altogether, and that those who argued for its continuation (with significant modifications) did so primarily because the study had already begun.

A few examples will illustrate the nature of the committee's objections to the WHI protocol. Thus, on the questionable hypothesis that a reduction in dietary fat will lead to a reduced incidence of breast cancer, the committee cites a number of large-scale studies that either refute or fail to find a direct correlation (see *Nature* 359, 760; 1992) and goes on to challenge NIH's rosy view that WHI will produce different results, especially in a short time.

The proposed protocol assumes a linear halving of the risk over five years... the biology of breast cancer would seem to make this an optimistic projection. If diet does have an effect on breast cancer, but the lag time for halving the risk is, for example, 20 years, then the currently proposed project has very little chance to detect an effect.

Why did NIH not think of that?

The NIH protocol is also criticized for the factorial design that simultaneously assigns some participants to two arms of the trial: diet modification and hormone replacement therapy. The committee argues on practical grounds that the chances of getting a woman to stick to a low-fat diet for 14 years while also staying on hormone therapy for the duration are slim. It adds that, even in the best case, the number of women in the overlap group (approximately 16 per cent) does not provide "adequate statistical power" to assess the interactions between the two regimens.

The proposals not to categorize women by known risk also provoke the committee's concern. Thus it argues that if risk factors, such as a family history of breast cancer, are not taken into account, the numbers of women in various of the study arms are too small to produce statistically meaningful data. Furthermore, the lack of information about risk factors would lead women to participate in a trial that could be hazardous to their health. There is, for example, nothing in the consent literature to warn a woman at risk from breast cancer that her risk might increase if, as one of the two-for-one participants, she agreed to take hormone therapy while also

restricting dietary fat.

On osteoporosis, the critics say that the NIH protocol will not allow useful distinctions to be made between calcium, vitamin D and the two in combination if, indeed, the trial data show that women in the study have fewer bone fractures. Similarly, the samples required by the protocols for assessing disease risk in minority groups (post-menopausal black and Hispanic women) are said to be too small to produce meaningful data. And the litany of criticisms goes on.

How can NIH, widely regarded as the fountainhead of the good design of trials and experiments, have perpetrated such a faulty design as that of WHI? NIH's intramural scientists, who would have responded to Healy's original call, have always been well regarded. But although there is always room for disagreement about the fine points of study designs and the interpretation of the data they yield, it is unusual for a review committee (whose own work is further subjected to the review board of the National Academy of Sciences, of which the IOM is a branch) to publish such a thorough condemnation of a study designed by fellow professionals.

A majority of the committee persuaded the remainder that WHI should be allowed to continue if NIH modifies its premises and corrects its statistical procedures, but the committee also acknowledges that politics, not science, has dissuaded it from outright cancellation of the study that Healy wrought.

In reality, committee members were besieged by women's health groups who said that cancelling the study would be a sign that the NIH does not care about women. Data indicating that approximately 14 per cent of NIH money is spent on diseases that affect women, compared with 80 per cent for research on diseases that afflict women and men equally, were sidelined by NIH officials seeking arguments in favour of a conspicuous Women's Health Initiative. But they have done women a disservice by backing a hastily put-together programme that is so vulnerable to informed criticism.

The NIH leadership has not distinguished itself by swaying with the political winds over the WHI. Even worse, it has done nothing to further the entirely valid need for more well-designed, well-funded studies of diseases to which women are particularly vulnerable. But WHI is just barely under way. Is it really too late to call a halt and, with science rather than politics as the motivation, start again?

Barbara J. Culliton

Heat-proof proteins

Norbert Hampp

PROTEINS are being exploited increasingly for technological ends, in biosensors and bioreactors for example. But some of these applications require high (more than 100 °C) operating temperatures, under which conditions many proteins are likely to denature. A striking exception is bacteriorhodopsin, as Shen *et al.* report on page 48 of this issue¹. They found that in dry films bacteriorhodopsin was structurally stable up to a temperature of 140 °C. This is an astonishing result and an encouraging one for a new aspect of bacteriorhodopsin research — its use in optical information processing.

Approaching San Francisco by air, one obtains a spectacular view of the intensely purple basins of the salt works in the bay area. In the 1970s it was shown that the colour was largely due to a retinal-containing protein which could be isolated from bacteria living in the harsh conditions of a saturated solution of common salt, with little dissolved oxygen². Survival in this ecological niche is possible because *Halobacterium salinarum* (formerly *Halobacterium halobium*) developed during evolution a set of proteins that enables it to use sunlight directly as an energy source. The key protein in the halobacterial photosynthesis is bacteriorhodopsin. Several tens of thousands of bacteriorhodopsin molecules form a two-dimensional crystalline lattice in the cell membrane over a region typically 500–1,000 nm in diameter and just 5 nm thick, the length of a single bacteriorhodopsin molecule. From its colour, this membrane fragment is called a 'purple membrane' and may be considered as the biological solar cell of the halobacterium. Each single bacteriorhodopsin molecule in the purple membrane acts as a light-driven proton pump. After absorption of a photon, bacteriorhodopsin transports a proton from the inside of the cell to the outer medium and thereby converts light energy into chemical energy which enables the halobacterial cell to survive.

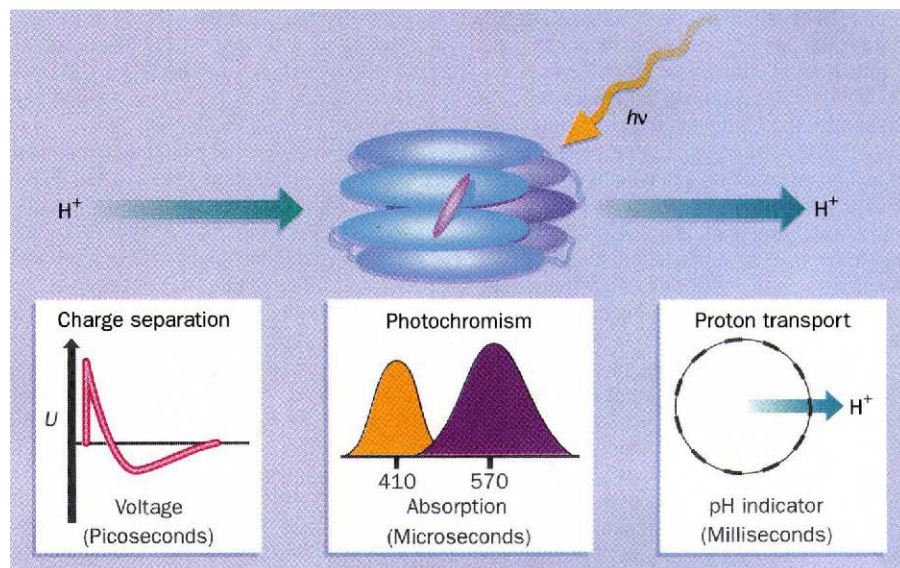
When we talk of bacteriorhodopsin's inertness towards chemical and photochemical degradation and its remarkable longevity, we should qualify this: it is the purple membrane form that we mean. Purple membranes contain about 10 lipid molecules per bacteriorhodopsin acting like a glue for the bacteriorhodopsin trimers. Solubilization of the membrane and analysis of the isolated trimeric or monomeric bacteriorhodopsin indicates that the intact membrane is crucial for the stability of the material. Shen *et al.*¹ have identified another key aspect, the absence of water, in maintaining stability above 100 °C, extending the utility of purple

membranes at the high temperatures demanded for technical applications.

Bacteriorhodopsin combines three functions that in principle are technically useful (see figure)³. First, it is a light-driven proton pump, so might be exploited to convert sunlight into chemical or electrical energy. Second, it has photoelectric properties due to the initial

have already proved useful in optical information processing, for example in holographic pattern recognition⁹. The optical properties of dry bacteriorhodopsin films might produce different technical applications¹⁰. Thermal denaturing has never been a problem, but nobody has tried to push dry bacteriorhodopsin films to the limits. The results of Shen and co-workers place them well into a temperature range that is adverse for most synthetic photochromic materials, especially in the presence of oxygen and light.

With luck, their work will encourage



Bacteriorhodopsin consists of a single strand of 248 amino acids arranged as seven α -helices (blue) and a covalently attached retinylidene residue (purple). Absorption of a photon leads to the transport of a proton through the proton pore of bacteriorhodopsin. This process divides into (1) an initial isomerization and charge separation on the picosecond timescale, (2) a reversible colour change on the microsecond timescale and finally (3) the proton transport which can be observed by means of pH indicators. A complete proton transport cycle of bacteriorhodopsin takes about 10 milliseconds. Each of these basic phenomena is of potential technical use.

charge separation in the molecule after absorption of a photon. Third, it is a photochromic protein: absorption of light leads to a reversible colour change from purple to yellow. It can be switched back and forth with blue and yellow light more often than any other known chemically synthesized photochromic compound. This and the fact that bacteriorhodopsin uses light very efficiently (just 1–2 photons are needed to switch a single molecule from purple to yellow) make it attractive for optical information processing.

Two decades of interdisciplinary research have pinned down the structure⁴ and molecular function^{5,6} of bacteriorhodopsin fairly precisely. Deliberately mutated bacteriorhodopsin molecules were first obtained in *Escherichia coli*⁷. With the development of a halobacterial transformation system⁸ it has become possible to obtain the mutated bacteriorhodopsin molecules in the stable purple membrane form and in bulk. Dry films of some functionally modified bacteriorhodopsins

others to transform 'well characterized' biological macromolecules into the solid or glassy state (not, in general, as easy as with purple membrane) and to investigate the conformational stability of the resulting material at higher temperatures. □

Norbert Hampp is at the Institute for Physical Chemistry, University of Munich, Sophienstrasse 11, D-80333 München, Germany.

- Shen, Y. *et al.* *Nature* **366**, 48–50 (1993).
- Oesterhelt, D. & Stoekenius, W. *Nature new Biol.* **233**, 149–152 (1971).
- Oesterhelt, D. *et al.* *Q. Rev. Biophys.* **24**, 425–478 (1991).
- Henderson, R. *et al.* *J. molec. Biol.* **213**, 899–929 (1990).
- Birge, R. R. A. *Rev. phys. Chem.* **41**, 683–733 (1990).
- Zimanyi, L. *et al.* *Biochemistry* **31**, 8535–8543 (1992).
- Braiman, M. S. *et al.* *J. biol. Chem.* **262**, 9271–9276 (1987).
- Ni, B. F. *et al.* *Gene* **90**, 169–172 (1990).
- Thoma, R. & Hampp, N. *Opt. Lett.* **17**, 1158–1160 (1992).
- Hampp, N. *et al.* *Mat. Res. Soc. Bull.* **17**, 56–60 (1992).

Hopes for hyperdeformation

W. R. Phillips

THINK of the atomic nucleus, and a picture of a roughly spherical conglomeration of protons and neutrons probably springs to mind. But 'superdeformed' nuclei, which are elongated into roughly ellipsoidal shapes with major axes about twice the length of their minor axes, are known to occur in some actinides^{1,2} at low spin and in some rare earths³ and nuclei near lead⁴ at high spin. Now, experiments by Galindo-Uribarri and colleagues indicate that some rapidly rotating nuclei may be still more severely elongated, with major and minor axes in a ratio near three to one⁵. These nuclei, predicted by some calculations, would be 'hyperdeformed'.

The shapes of objects depend on their internal construction, which can be examined by determining shapes under different conditions and unravelling the role the objects' constituents play in producing them. This is as true of the smallest discrete objects in nature, such as atomic nuclei, as it is of the largest, such as rotating stars and galaxies. Nonspherical microscopic objects, diatomic molecules, for example, exhibit quantum states which correspond to rotations. Rotation at higher and higher speeds produces a sequence of states with energies that depend on the quantized angular momentum in the rotation and on the moment of inertia associated with it. Observa-

tions of the photons produced by decays between levels in such a rotational band determine the moment of inertia and then, indirectly, the shape of the rotating object. This forms the basis of the experiment to search for hyperdeformed nuclei.

Compound nuclei can be made in the fusion of a heavy-ion beam with nuclei in a target. The compound nuclei are hot and highly excited, and in a distribution of spin states, sometimes up to spins as high as $80\hbar$ or more, where $\hbar$ is the quantum unit of angular momentum. The hot nuclei cool down, usually by emitting neutrons, and leave lighter nuclei which de-excite to their ground state by γ -ray emission. The first few γ -ray decays usually lead to levels of high spin in rotational bands, from where the ground states are reached by more γ -ray transitions down the bands.

The γ -ray energies contain the signatures of the moments of inertia. Making the simplifying assumption that there is only one rotational band with a moment of inertia I_m , the energy of a band level with spin quantum number I is given by $E(I) = \hbar^2 I(I+1)/2I_m$, and the energy of a transition to a lower level is proportional to the spin of the upper level and inversely proportional to the moment of inertia.

The transitions from a range of levels are shown schematically in part *a* of Fig. 1. Plotted against energy, they form a picket fence, with the difference in energies of adjacent lines constant and inversely proportional to I_m . The transitions in the decay of a compound nucleus all occur essentially simultaneously, and experiments can be designed to pick out coincidences between any two γ -rays in the decays. The coincidence events can then be placed in a two-dimensional array (matrix), one side of the matrix corresponding to the energy of one γ -ray and the second side corresponding to the energy of the other (part *b* of Fig. 1). As the lines in the decays down the rotational band all have different energies, and one transition cannot produce two γ -rays, the diagonal of this matrix is empty. Instead, the matrix should contain events dotted along equidistant parallel lines either side of the diagonal ('ridges'; for example *xx'* and *yy'* in part *b* of

Fig. 1). The separation of the ridges contains the required information on the moment of inertia.

In an actual experiment this simple pattern of ridge structures would be greatly modified. The 'events' could spring up all over the matrix because of incorrect energy signals due to detector imperfections, or false coincidences arising from transitions between many dif-

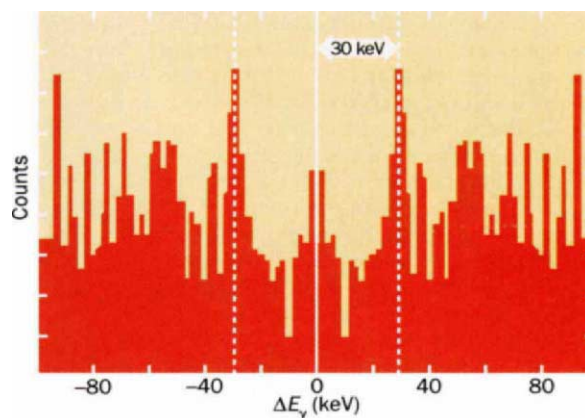


FIG. 2 Distribution of events within a region of the experimentally determined coincidence matrix, projected onto an axis perpendicular to the diagonal (parallel to the line *XX'* in Fig. 1). The abscissa gives the differences in energies of the coincident γ -rays, and the events are for γ -ray energies near 1,440 keV. The pattern is consistent with ridges 30 keV either side of the diagonal, and is symmetrical about the diagonal because of the way the data are processed.

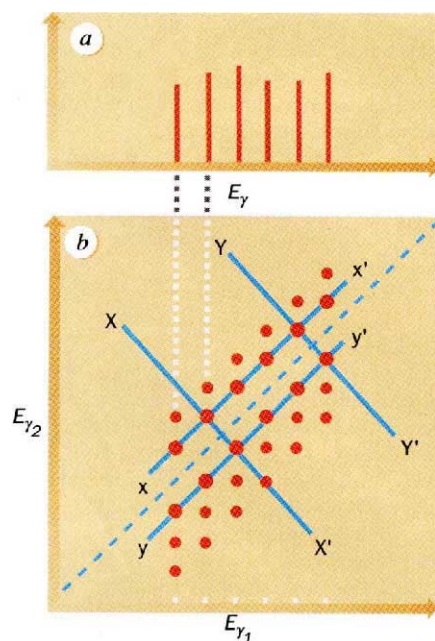


FIG. 1 *a*, Uniform spacings of the energies of transitions between levels in an ideal rotational band. *b*, 'Matrix' constructed with the energy of one transition between excited states on one side and the energy of a second, coincident transition on the other. They would lie on parallel lines such as *xx'* and *yy'*.

ferent levels rather than those in a single rotational band. In addition, events could arise from more than one nucleus, several bands could contribute to the decay paths in a single nucleus, and even in a single band the energy spacings do not precisely follow the simple formula quoted above.

All these factors make the ridge structures in the matrix hard to see. However, under some circumstances and over limited ranges of γ -ray energies, ridges may be observed and provide the required information on a rotational band.

Hyperdeformed states are expected to occur at high spin only. It is not intuitively obvious, but electrostatic forces between protons in the compound nucleus impede charged-particle emission in the cooling process. If a nucleus is elongated, however, the barrier against proton emission is reduced at the tips, and some cooling may take place via protons. So decays down hyperdeformed bands may be enhanced if γ -ray coincidence events are selected only when they accompany proton emission. Galindo-Uribarri and colleagues⁵ used this trick to increase experimental sensitivity in their experiment, in which a beam of ^{37}Cl ions (from the accelerator at the Chalk River Laboratories, Canada) bombarded ^{120}Sn nuclei to make a ^{157}Ho compound nucleus; the loss of one proton (and an unspecified

number of neutrons) from holmium produces dysprosium, Dy. Coincident γ -rays were observed in an array of 20 germanium detectors. A matrix of coincident γ -rays in dysprosium nuclei was obtained, and the kinematics of the reaction suggested that the most probable dysprosium isotopes being studied were ^{152}Dy and ^{153}Dy .

Simple structures resembling those in Fig. 1 should show up when the events are projected onto an axis perpendicular to the diagonal. For example, the projection on the perpendicular axis of events between lines XX' and YY' would show a dip at the position of the diagonal, and peaks where the perpendicular crosses the ridges. The first ridges, either side of the middle, would be most prominent with successive ridges more and more smeared out because of non-ideal spacings of the band members. Superdeformed bands are known⁶ to occur in ^{153}Dy and to show up as ridges 48 keV either side of the diagonal for γ -ray energies near 1,300 keV. The appropriate projection of the data in the Chalk River matrix did indeed show such ridges.

But that was not all. A projection of the data near γ -ray energies of 1,440 keV gave the pattern shown in Fig. 2; this is as expected for new ridges just 30 keV either side of the diagonal. The distance suggests a rotational band with a moment of inertia much larger than that of the superdeformed bands in dysprosium nuclei. If a simple model of a rotating rigid ellipsoid is used to translate the moment of inertia into deformation, the major-to-minor-axis ratio deduced is near three to one — hyperdeformation.

It is too early for the cautious to be sure that this interpretation is correct. The discrete transitions within the hyperdeformed band have not yet been clearly observed, and the pathway from the observation of ridge structures to the deduction of band deformations is strewn with pitfalls. Yet the first indications⁷ of superdeformation, later substantiated³, came from experiments much like this. Those designing further experiments to determine the shapes of nuclei under extreme conditions should be encouraged by the results. □

W. R. Phillips is in the Department of Physics, Nuclear Group, Schuster Laboratory, University of Manchester, Manchester M13 9PL, UK.

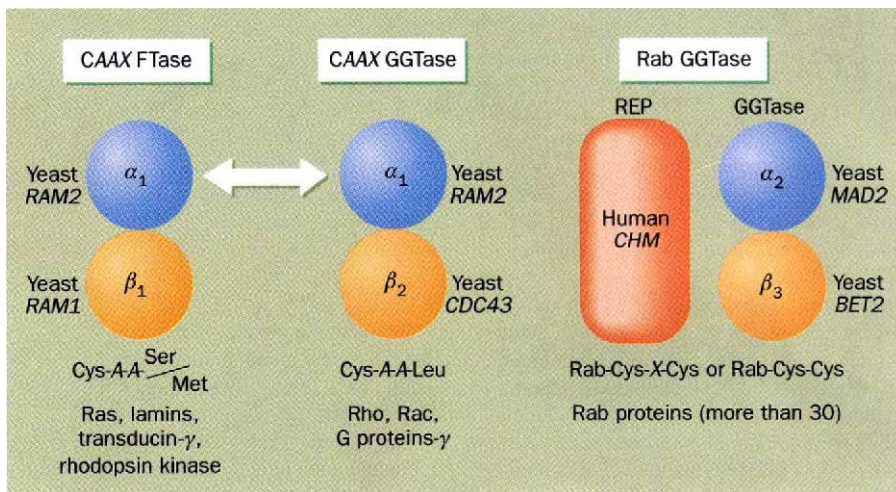
Mad Bet for Rab

Michael S. Brown and Joseph L. Goldstein

GENETICISTS and biochemists often argue about whose world-view is more powerful. The answer, of course, is neither — both would be lost without the molecular biologists and their sequencing machines. The synergism of these disciplines is illustrated in the brand-new field of protein prenylation, in which mutant phenotypes have been assigned to all six prenyltransferase subunits by sequence comparison

Rabs regulate fusion of vesicles during secretion and endocytosis. Prenylation is essential for the function of all these proteins^{3–5}.

Ras proteins, like mating factors, undergo farnesylation at the cysteine (C) of a carboxy-terminal CAA X box (A is an aliphatic residue and X is usually methionine or serine), followed by proteolysis and methylation. Most Rho/Rac



Three heterodimeric prenyltransferases and an escort protein (REP) have been characterized biochemically. Mutations in each subunit have been described in yeast or humans. The α - and β -subunits are distinguished by numerical subscripts, and the preferred substrates for prenyltransferase activity are indicated at the bottom.

(see figure). The letters by Li *et al.*¹ and Jiang *et al.*² on pages 82 and 84 of this issue complete the mutational catalogue by nailing down the involvement of the *MAD2* and *BET2* genes of yeast. These findings have immediate implications for the control of membrane movement and cell growth in eukaryotes.

The biochemical story begins in fungi with the discovery that mating proteins possess prenyl groups, polyunsaturated hydrocarbons composed of repeating five-carbon isoprenes. Farnesyl, the prenyl group of mating factors, contains three isoprenes (15 carbons) and is linked by a thioether bond to a cysteine at the fourth position from the carboxy terminus. Thereafter, the three terminal residues are removed, and the cysteine is rendered hydrophobic by carboxy-methylation^{3,4}.

In animals, most but not all prenylated proteins are small GTP-binding proteins (relative molecular mass 20–30K) whose regulatory actions are initiated by binding GTP and terminated by slow GTP hydrolysis. These proteins belong to three families: Ras, Rho/Rac and Rab⁵. Ras proteins regulate growth by transmitting signals from the plasma membrane to the nucleus; Rho and Rac regulate cytoskeleton-membrane interactions; and

proteins are processed similarly except that X is leucine, and geranylgeranyl (GG) (four isoprenes; 20 carbons) replaces farnesyl. Rabs lack CAA X boxes. Most terminate in closely spaced cysteines, either CC or CXC. Rab3A, for example, terminates in the sequence cysteine-alanine-cysteine. Both cysteines are geranylgeranylated, and the terminal cysteine is methylated⁶.

One farnesyltransferase (FTase) and two GG transferases (GGTases) have been purified from animals (see figure). The prototype is CAA X FTase, an $\alpha\beta$ heterodimer, which forms a carrier complex with farnesyl pyrophosphate^{7,8}. The β -subunit binds Zn^{2+} and the CAA X box, but is inactive without the α -subunit⁹. CAA X GGTase (or GGTase I), also an $\alpha\beta$ heterodimer, recognizes leucine-terminated CAA X boxes in Rac/Rho proteins. The α -subunit is identical to the α -subunit of FTase¹⁰.

The third prenyltransferase is Rab GGTase (or GGTase II) (see figure). Whereas the other two enzymes recognize CAA X boxes, Rab GGTase recognizes more extensive, and still unidentified, internal sequences common to Rabs. Recognition is mediated by an accessory protein, originally called component A and

- Polikanov, S. M. *et al.* *Sov. Phys. JETP*, **15**, 1016–1021 (1962).
- Strutinsky, V. M. *Nucl. Phys. A*, **95**, 420–442 (1967); **A122**, 1–33 (1968).
- Twin, P. J. *et al.* *Phys. Rev. Lett.*, **57**, 811–814 (1985).
- Moore, E. F. *et al.* *Phys. Rev. Lett.*, **63**, 360–363 (1989).
- Galindo-Uribarri, A. *et al.* *Phys. Rev. Lett.*, **71**, 231–234 (1993).
- Johansson, J. K. *et al.* *Phys. Rev. Lett.*, **63**, 2200–2203 (1989).
- Nyako, B. M. *et al.* *Phys. Rev. Lett.*, **52**, 507–510 (1984).

renamed rab escort protein (REP). This binds Rabs and presents them to Rab GGTase (originally called component B)^{11,12}, a heterodimer whose α - and β -subunits resemble those of CAAX FTase¹³.

The genetic story begins in yeast with the identification of *RAM1* (or *DPR1*), a suppressor of overactive mutant Ras^{3,4}. *ram1* mutants fail to farnesylate α -mating factor as well as Ras. Mutations at another locus, *RAM2*, exhibit a similar phenotype^{3,4}. *RAM1* and *RAM2* encode the β - and α -subunits of CAAX FTase, respectively^{3,8}. *ram2* mutants are deficient in CAAX GGTase activity as well as FTase activity, indicating that in yeast, as in animals, the α -subunit is shared¹⁴.

Yeast with mutations in the *CDC43* (or *CAL1*) gene have defects in the cell cycle and lack CAAX GGTase activity. The gene product resembles the β -subunit of CAAX GGTase³.

The first mutation affecting Rab GGTase was discovered not in yeast, but in humans with choroideremia (CHM), an X-linked retinal degeneration. The human *CHM* gene, isolated by positional cloning^{15,16}, encodes a protein that is 87 per cent the same in sequence as rat REP-1^{11,12}. Choroideremic lymphoblasts are deficient in Rab GGTase activity and are corrected *in vitro* by rat REP-1¹⁷. It is surprising that deletions of REP-1 lead only to retinal degeneration — after all, geranylgeranylated Rabs should be crucial for secretion and exocytosis in all cells. This paradox may be explained by the *CHML* (choroideremia-like) gene on chromosome 1¹⁸, in that its product, REP-2, supports geranylgeranylation of most Rab proteins¹⁹, and may substitute for

REP-1 in tissues other than retina.

Li *et al.*¹ and Jiang *et al.*² have now identified mutations in the last two prenyltransferase subunits, Rab GGTase α and β , encoded by yeast genes *MAD2* and *BET2*, respectively. The Mad2 protein is a diminutive version of rat Rab GGTase α (290 against 567 amino acids)^{13,20}. Bet2 is 51 per cent the same as rat Rab GGTase β ¹³. The *bet2* secretory mutant fails to geranylgeranylate two Rab proteins, Ypt1 and Sec4, which are required for vesicular transport. In a welcome marriage of genetics and biochemistry, Jiang *et al.* demonstrate that Bet2 protein forms an active complex with Mad2 protein, and that GG transfer is stimulated by a factor in yeast extracts, presumably yeast REP.

The *MAD2* gene is puzzling because two of its mutant alleles give drastically different phenotypes. Cells with the *mad2-1* mutation fail to arrest mitosis when microtubular polymerization is slowed by inhibitors²¹. Cells with a *ts* allele

(*mad2-2^{ts}*) fail to geranylgeranylate Ypt1, Sec4 and other proteins at the nonpermissive temperature, and are secretion-defective like *bet2* cells. Strangely, *mad2-2^{ts}* cells arrest mitosis normally¹. Conversely, *mad2-1* cells have no consistent defect in geranylgeranylation¹. They may produce a defective α -subunit that supports geranylgeranylation of Ypt1 and Sec4, but not another Rab that regulates mitotic arrest; alternatively, the α -subunit may perform an unrelated function in mitosis that is selectively abolished in *mad2-1* cells. Once again, these inconsistencies between genetics and biochemistry will be solved by enlisting the aid of molecular biologists and their all-powerful sequencing machines. □

Michael S. Brown and Joseph L. Goldstein are in the Department of Molecular Genetics, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235, USA.

DIABETES

Spotlight on a neuronal enzyme

Michele Solimena and Pietro De Camilli

INSULIN-DEPENDENT diabetes mellitus (IDDM) results from autoimmune destruction of the insulin-secreting β -cells of the pancreas. Ever since the discovery in IDDM patients of autoantibodies directed against islet cells and of lymphocytic infiltrates in islets (insulinitis), a major goal has been the identification of the autoantigen(s) that drive(s) the autoimmune response. Among the autoantibodies known to be associated with IDDM, those directed against the enzyme glutamic acid decarboxylase (GAD) are the ones that appear earliest and are present in the largest number of patients. GAD has been receiving increasing attention as a putative key autoantigen, and now two reports on pages 69 and 72 of this issue^{1,2} direct the spotlight onto GAD as a central player in the early events leading to IDDM in NOD mice, a mouse strain that spontaneously develops the disease and is used as an animal model of IDDM.

GAD synthesizes γ -aminobutyric acid (GABA), the principal inhibitory neurotransmitter in the brain and a putative paracrine signal molecule in pancreatic islets³. The enzyme, which has a restricted tissue distribution and is highly expressed in the cytoplasm of GABA-secreting neurons and pancreatic β -cells, is represented by two isoforms, referred to as GAD65 and GAD67 according to their relative molecular masses⁴. The two proteins are the products of two different genes and differ substantially only in their amino-terminal region (see figure overleaf), which in GAD65 contains the

information responsible for its targeting to GABA-containing secretory vesicles.

GAD was first identified as an autoantigen in stiff-man syndrome, a rare disorder of the central nervous system characterized by progressive rigidity of the body musculature with painful spasms⁵. High-titre autoantibodies directed against GAD were detected in the serum and cerebrospinal fluid of 60 per cent of patients with the syndrome. Strikingly, many of these patients had late-onset IDDM and a few IDDM patients without the syndrome, but with the same antibodies, were identified^{5,6}. These findings suggested a possible link of GAD autoimmunity not only with stiff-man syndrome, but also with at least some cases of IDDM.

A β -cell autoantigen of relative molecular mass 64,000 (64K), first described in the early 1980s, had independently become the focus of increasing attention in the field of IDDM research⁷. Antibodies directed against this protein, originally thought to be β -cell specific, were found in most newly diagnosed diabetic patients and in prediabetic patients long before the onset of clinical symptoms, suggesting that this antigen could be important in the early stages of the disease. These antibodies were also found in animal models of IDDM, including the NOD mouse^{8,9}. Technical limitations affected the purification of this protein from pancreatic islets and delayed its further characterization as an autoantigen.

The convergence of the two lines of

1. Li, R., Havel, C., Watson, J. A. & Murray, A. W. *Nature* **366**, 82–84 (1993).
2. Jiang, Y., Rossi, G. & Ferro-Novick, S. *Nature* **366**, 84–86 (1993).
3. Schafer, W. R. & Rine, J. A. *Rev. Genet.* **30**, 209–237 (1992).
4. Gibbs, J. B. *Cell* **65**, 1–4 (1991).
5. Takai, Y., Kaibuchi, K., Kikuchi, A. & Kawata, M. *Int. Rev. Cytol.* **133**, 187–230 (1992).
6. Glomset, J. A., Gelb, M. H. & Farnsworth, C. C. *Curr. Opin. Lipidol.* **2**, 118–124 (1991).
7. Reiss, Y., Goldstein, J. L., Seabra, M. C., Casey, P. J. & Brown, M. S. *Cell* **62**, 81–88 (1990).
8. Chen, W.-J., Andres, D. A., Goldstein, J. L., Russell, D. W. & Brown, M. S. *Cell* **66**, 327–334 (1991).
9. Reiss, Y., Brown, M. S. & Goldstein, J. L. *J. Biol. Chem.* **267**, 6403–6408 (1992).
10. Seabra, M. C., Reiss, Y., Casey, P. J., Brown, M. S. & Goldstein, J. L. *Cell* **65**, 429–434 (1991).
11. Seabra, M. C., Brown, M. S., Slaughter, C. A., Südhof, T. C. & Goldstein, J. L. *Cell* **70**, 1049–1057 (1992).
12. Andres, D. A. *et al.* *Cell* **73**, 1091–1099 (1993).
13. Armstrong, S. A., Seabra, M. C., Südhof, T. C., Goldstein, J. L. & Brown, M. S. *J. Biol. Chem.* **268**, 12221 (1993).
14. Moore, S. L. *et al.* *J. Biol. Chem.* **266**, 14603 (1991).
15. Cremers, F. P. M. *et al.* *Nature* **347**, 674–677 (1990).
16. Merry, D. E., Jänne, P. A., Landers, J. E., Lewis, R. A. & Nussbaum, R. L. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2135–2139 (1992).
17. Seabra, M. C., Brown, M. S. & Goldstein, J. L. *Science* **259**, 377–381 (1993).
18. Cremers, F. P. M. *et al.* *Hum. molec. Genet.* **1**, 71 (1992).
19. Cremers, F. P. M., Armstrong, S. A., Seabra, M. C., Brown, M. S. & Goldstein, J. L. *J. Biol. Chem.* (in the press).
20. Boguski, M. S., Murray, A. W. & Powers, S. *New Biol.* **4**, 408–411 (1992).
21. Li, R. & Murray, A. W. *Cell* **66**, 519–531 (1991).

research with the identification of the 64K antigen of IDDM as GAD¹⁰ sparked a burst of investigation of this protein. Its localization in β -cells of pancreatic islets made it an attractive candidate participant in an autoimmune process known to spare other islet cells. On the other hand, the cytoplasmic localization of this antigen, as well as the solid evidence indicating that

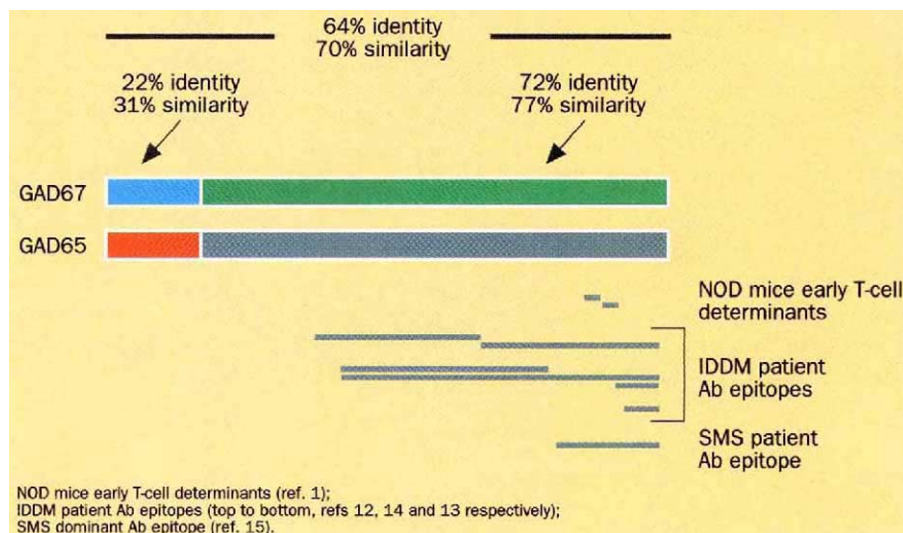
for human IDDM. The progressive intramolecular spreading of T-cell reactivity is reminiscent of what has been found for T-cell responses to myelin basic protein in experimental allergic encephalomyelitis¹¹.

The two studies also show that intravenous¹ or intrathymic² injection of GAD65 in three-week-old NOD mice

destruction inevitably occur.

Third, the T-cell responses to GAD indicate a primary autoantigenic role for the C-terminal region of GAD65, a region where GAD65 and GAD67 are very similar (see figure); this region appears to be crucial for autoantibody reactivity as well. Many patient autoantibodies recognize epitopes that are located in or include (conformational epitopes) the C-terminal region of GAD65 (refs 12–15), but do not recognize the very similar C-terminal region of GAD67 (refs 14,15). This apparent paradox may hide an explanation for autoimmunity to GAD. During ontogenetic development, GAD67 appears to be expressed before GAD65 (M. Erlander, personal communication). Early presentation of GAD67 may eliminate T cells strongly reactive against GAD67 but preserve those that are weakly reactive. When GAD65 is expressed, the weakly reactive cell population, within a given genetic background, may then react strongly with GAD65 peptides that are similar to, yet slightly different from, the corresponding peptides of GAD67. This process, reminiscent of the induction of experimental autoimmunity to a given antigen by injection of a corresponding heterologous protein¹⁶, would eventually lead to loss of tolerance also for GAD67. Alternatively, some unique biochemical properties of GAD65, such as its hydrophobicity resulting from post-translational modifications¹⁴, may be responsible for an easier loss of tolerance to this GAD isoform.

Fourth, and most interesting, is the finding that the humoral response in GAD65-injected mice persists, what Tisch *et al.* define as a "split tolerance"². This is directly relevant to the observation that in humans there is an inverse relation between humoral (Th2 T-cell-mediated) and cellular (Th1 T-cell-mediated) autoimmunity to GAD in subjects at risk of IDDM¹⁷. It was also shown that a strong



Domain structure of human GAD and autoantigenic regions of GAD65. Percentages of identity and similarities between human GAD65 and GAD67 were calculated using Genework software version 2.2 and the GAD sequences published in ref. 19. Light green bars drawn below the GAD65 molecule depict regions of GAD65 that react with NOD mouse T cells or human autoantibodies. Data from Richter *et al.*¹⁴ refer to monoclonal antibodies. The T-cell peptides shown are only the ones involved in the initial response in the NOD mouse; subsequently T-cell responses spread to other peptides. Only the dominant epitope of SMS is shown.

IDDM is a T-cell-mediated disease, argued against a direct role of anti-GAD autoantibodies in β -cell destruction.

The two letters in this issue^{1,2} provide evidence that loss of tolerance to GAD65 is an early and necessary step towards the development of diabetes in NOD mice. They show that a T-cell response to GAD65 develops at four weeks of age, at the same time as the onset of insulinitis, and before T-cell responses arise to a panel of other IDDM autoantigens. At the same age, autoantibodies against GAD65 and GAD67 also appear². The T-cell response is accounted for by CD4⁺Th1 lymphocytes, consistent with a presentation of GAD peptides by class II major histocompatibility complex molecules and with a primary involvement of a subset of T-helper cells (CD4⁺Th1 cells) that sustain primarily cell-mediated immune reactions. As shown in one of the two reports, T-cell reactivity is initially restricted to the C-terminal regions of GAD65 (see figure), but later spreads to other parts of GAD65, including a region with sequence similarity to a protein of Coxsackie virus¹. This chain of events speaks against the hypothesis that molecular mimicry between Coxsackie virus and GAD65 triggers the autoimmune response in NOD mice, although it leaves the question open

markedly reduces T-cell proliferative responses both to GAD and to other IDDM autoantigens, and prevents both insulinitis and diabetes^{1,2}. By contrast, injection of other IDDM autoantigens induces tolerance to the specific injected protein, but not to other autoantigens, including GAD, and does not prevent insulinitis and diabetes. The production of interferon- γ in antigen-stimulated spleen cell cultures from GAD65-tolerant mice is significantly reduced, indicating that tolerance may result from suppression of GAD65-responsive Th1 T cells. This effect is not paralleled by a corresponding reduction of the humoral response to GAD and to the other IDDM autoantigens, suggesting a persistent response of CD4⁺Th2 T cells². These are the subset of T-helper cells that primarily mediate humoral immune responses.

These findings support several interesting conclusions. First, an autoantigen originally identified as a target of humoral autoimmunity may also be critical in T-cell-mediated autoimmunity. Second, regardless of whether GAD is the first antigen for which tolerance is lost, the appearance of a CD4⁺Th1 T-cell response to the C-terminal region of GAD65 seems to be a turning point, at least in NOD mice, after which insulinitis and β -cell

1. Kaufman, D. L. *et al.* *Nature* **366**, 69–72 (1993).
2. Tisch, R. *et al.* *Nature* **366**, 72–75 (1993).
3. Reetz, A. *et al.* *EMBO J.* **10**, 1275–1284 (1991).
4. Erlander, M. G., Tillakaratne, N. J. K., Feldblum, S., Patel, N. & Tobin, A. J. *Neuron* **7**, 91–100 (1991).
5. Solimena, M. *et al.* *New Engl. J. Med.* **318**, 1012–1020 (1988).
6. Solimena, M., Folli, F., Aparisi, R., Pozza, G. & De Camilli, P. *New Engl. J. Med.* **322**, 1555–1560 (1990).
7. Bækkeskov, S. *et al.* *Nature* **298**, 167–169 (1982).
8. Sigurdsson, E. & Bækkeskov, S. *Curr. Top. Microbiol. Immun.* **184**, 143–167 (1990).
9. Atkinson, M. A., Maclaren, N. K., Scharp, D. W., Lacey, P. E. & Riley, W. J. *Lancet* **335**, 1357–1360 (1990).
10. Bækkeskov, S. *et al.* *Nature* **347**, 151–156 (1990).
11. Lehmann, P. V., Forsthuber, T., Miller, A. & Sercarz, E. E. *Nature* **358**, 155–157 (1992).
12. Kaufman, D. L. *et al.* *J. Clin. Invest.* **89**, 283–292 (1992).
13. Mauch, L. *et al.* *Eur. J. Biochem.* **212**, 597–603 (1993).
14. Richter, W., Shi, Y. & Bækkeskov, S. *Proc. natn. Acad. Sci. U.S.A.* **90**, 2832–2836 (1993).
15. Butler, M. H., Solimena, M., Dirks, R., Hayday, A. & De Camilli, P. *J. exp. Med.* (in the press).
16. Mamula, M. J., Lin, R. H., Janeway, C. A. Jr & Hardin, J. A. *J. Immun.* **149**, 789–795 (1992).
17. Harrison, L. C. *et al.* *Lancet* **341**, 1365–1369 (1993).
18. Gianani, R. *et al.* *Diabetes* **41**, 347–353 (1992).
19. Bu, D. F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **89**, 2115–2119 (1992).

humoral response to GAD correlates with a slow progression to IDDM^{17,18}. Accordingly, GAD65 autoantibody titres are higher and more persistent in stiff-man patients, who develop IDDM when they are older^{5,6}, than in patients affected by the common juvenile form of IDDM^{17,18}. It would be of interest to know whether some of the GAD65-injected NOD mice, in which the humoral response to GAD is preserved, eventually develop a neurological disease resembling stiff-man syndrome. This would have obvious implications for the development of immunotherapies in humans. An animal model for this syndrome, in which the pathogenetic mechanisms are unclear, would be most useful in the further elucidation of

molecular events underlying GAD autoimmunity.

Beyond these speculations, the demonstration that injection of GAD65 can prevent spontaneous IDDM in a syngenic animal model of the disease raises high expectations. If future work shows that the findings of Kaufmann *et al.*¹ and Tisch *et al.*² are also applicable to humans, the manipulation of the immune response to GAD may help to alleviate some of the suffering caused by IDDM. □

Michele Solimena and Pietro De Camilli are in the Department of Cell Biology and Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

HYDROLOGICAL CYCLE

Drought in Greenland?

Kenneth C. Jezek

ICE sheets are nourished by the annual accumulation of snow precipitated from storm systems traversing across their surfaces. Annual accumulation serves to balance ice wastage from the boundary regions of the ice sheets and so is one of the two key processes that maintain the vast reservoirs of fresh water at present bound up in the ice sheets. Long records of annual accumulation are also decipherable from the physical and chemical properties of ice cores extracted from the depths of the ice sheets, which can be used to reveal annual variations in the Earth's climate over hundreds of thousands of years.

Information about the modern patterns of precipitation and accumulation is largely derived from direct measurements of accumulation on the ice-sheet surface. For instance, time-averaged descriptions of precipitation and accumulation over the Greenland Ice Sheet come from analyses by Ohmura and Reeh¹, based on accumulation measurements from nearly 250 snow pits and ice cores obtained between 1913 and 1989, together with precipitation observations from 35 coastal meteorological stations from 1931 to 1980. Bolzan and Strobel² analysed ice cores about 17 m deep collected over a 150-by-150-km area of north-central Greenland to infer the annually varying accumulation there over the period 1964–87. Now the modellers are catching up. In the *Journal of Climate*, David Bromwich and his co-workers³ provide interannual estimates of precipitation integrated over the entire Greenland Ice Sheet.

In their approach, variability in the location and intensity of storm tracks and the effects of ice-sheet topography on atmospheric flow are incorporated in a diagnostic numerical model of mid-

tropospheric circulation. Climatological data supplied by the National Meteorological Center are combined with Bolzan and Strobel's earlier 'point' measurements to let them describe precipitation in terms of short-wave vorticity advection, lower-tropospheric wind velocities, and the height of the ice sheet. Available climate data allow them to compute annual precipitation patterns for 1963–89.

Several of the results from this analysis are intriguing. For a start, Bromwich's team finds that there is a pronounced downward trend of about 15 per cent in modelled total annual precipitation over the 26-year period. In this time of suspected global warming, one might expect

an increase in precipitation because of the exponentially larger moisture-holding capacity of warmer air and the increasing exposure of ocean water from a generally retreating sea-ice cover. But temperatures in the Greenland region have in fact decreased over the past three decades (although as Bromwich and colleagues point out, temperature is only one factor associated with decreases in precipitation) and sea-ice extents have increased to the west of Greenland⁴. Sea ice is far more ephemeral than the ice sheet itself, more easily broken and dispersed; the further the sea ice extends around the fringes of the ice sheet, the more it will limit ocean surface evaporation and reduce that local contribution to precipitation. Decreasing precipitation also suggests that less water is being bound up in the ice sheet. Although it is not clear how the remaining atmospheric water is redistributed, Bromwich suggests that this deficit in snow accumulation could result in an increase of 0.3 millimetres per year in global sea level.

Details of how energy is transported from the Equator to the poles — in other words, the routes taken by storm systems — turn out to be crucial. Bromwich's team find strong (30 per cent) year-to-year variations in accumulation, which they interpret as caused by relatively slight changes in the North Atlantic storm track. Accumulation rates derived from ice cores also show strong fluctuations and variations on very short timescales; for example, an abrupt doubling in accumulation is seen at the end of the Younger Dryas cold snap⁵. Perhaps Bromwich's calculations hint at the cause for the ice-core observations: a moderate, sustained change in the North Atlantic storm track would have

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

A question of balance — mass lost from the Greenland Ice Sheet, in part through icebergs such as these, must be replenished by precipitation. (Courtesy of J. A. Dowdeswell.)

had a considerable effect on the accumulation rate over Greenland, without the need for dramatic shifts in the planetary (or hemispheric) atmospheric circulation. Admittedly it isn't clear why a change should have been sustained for thousands of years. Nonetheless, the results highlight the need to understand how far the ice-core observations reflect global as opposed to merely regional climate changes.

The annual variation also has implications for the interpretation of changes in ice-sheet elevation estimated by Zwally and others⁶ using data from spaceborne altimeters. Modelling ice dynamics serves to integrate the effects of short-term changes in snow accumulation on the long-term mass balance of the ice sheet through the relatively long response time of the ice sheet. But those short-term variations will still be evident in rapid fluctuations in surface elevation. The modelling results indicate a secondary rise in precipitation rates during the period from 1977 to 1985, just the period for which Zwally and others observe an increase in ice-sheet elevation. As the model results imply far less increase in elevation than was observed, though, the altimeter results must remain controversial⁷.

Bromwich and Robasky⁸ show elsewhere that the modelled precipitation trends for 1963–89 are spatially variable, with increases, particularly in northeastern Greenland, co-existing with the predominantly decreasing trends. Increasing modelled precipitation over northeastern Greenland is consistent with recent estimates of accumulation increases in this area for the 1979–87 period⁹, so it looks as if their calculation is along the right lines.

If we are lucky, similar work in the Antarctic will fill in the broader picture. Bromwich and Robasky⁸ say, on the basis of accumulation records and analyses of the atmospheric moisture budget, that precipitation rates have recently increased for the Antarctic Ice Sheet, the opposite effect to what is seen in Greenland.

As yet, we are only scratching the surface of the complex interactions between precipitation, air temperature and storm tracks over the ice sheets, any or all of which could be affected by global warming. □

Kenneth C. Jezek is at the Byrd Polar Research Center, Ohio State University, 1090 Carmack Road, Columbus, Ohio 43210, USA.

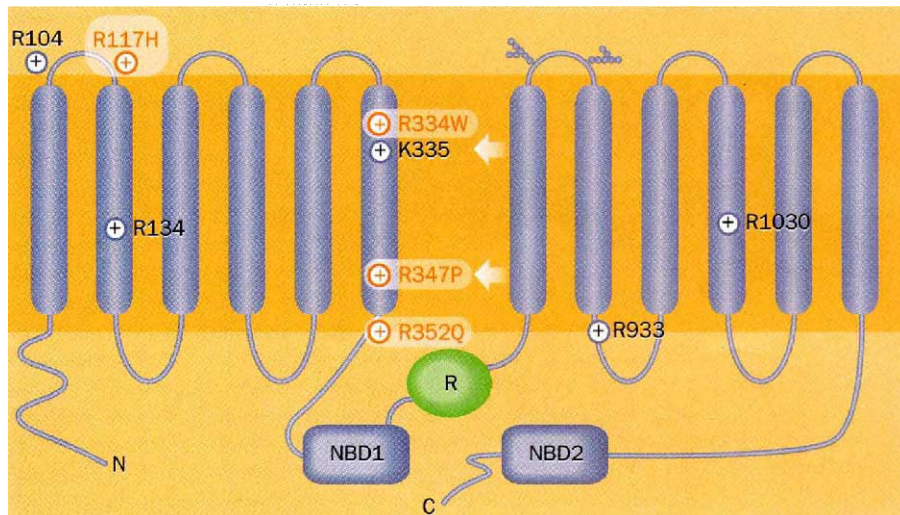
- Ohmura, A. & Reeh, N. J. *Glaciol.* **37**, 140–148 (1991).
- Bolzán, J. F. & Strobel, M. J. *Glaciol.* (in the press).
- Bromwich, D. H. *et al.* *J. Clim.* **6**, 1253–1268 (1993).
- Chapman, W. L. & Walsh, J. E. *Bull. Am. met. Soc.* **74**, 33–47 (1993).
- Alley, R. B. *et al.* *Nature* **362**, 527–529 (1993).
- Zwally, H. J. *et al.* *Science* **246**, 1587–1589 (1989).
- Van der Veen, C. J. *Clim. Change* **23**, 383–405 (1993).
- Bromwich, D. H. & Robasky, F. M. *Met. Atmos. Phys.* **51**, 259–274 (1993).
- Steffen, K. *et al.* *Met. Atmos. Phys.* **51**, 239–258 (1993).

Indictment of pore behaviour

Jeffrey J. Wine

THE elegant work published on page 79 of this issue¹ will be of considerable interest to biophysicists generally, as well as to the cystic fibrosis research community. Tabcharani *et al.* show that the cystic fibrosis transmembrane conductance regulator (CFTR), the chloride channel protein that is defective in cystic fibrosis, normally operates as a multi-ion pore. Moreover,

ion at a time, and the consequences for permeation properties can be dramatic. 'Anomalous mole-fraction behaviour' is one such consequence⁵. Tabcharani *et al.* discovered anomalous mole-fraction behaviour for CFTR conductance, and then used it to assay the effects of mutations on the permeation pathway of CFTR. With 154 mM Cl[−] on both sides of the mem-



Charged residues in or near the putative transmembrane domains of the CFTR protein. Arrows show residues that were mutated by Tabcharani *et al.*¹. Residues in red type are sites of mutations that cause CF (the letter preceding the number denotes the wild-type amino acid; that following, the mutant). R, regulatory phosphorylation domain. NBD, nucleotide-binding domain.

disease-causing mutations of a particular amino acid convert CFTR to a single-ion pore with reduced conductance.

Cystic fibrosis is a recessive genetic disease with a Caucasian carrier frequency of about 4 per cent. Its primary target is exocrine glands, and it is caused by mutations that disrupt CFTR². A vigorous consortium has identified hundreds of disease-causing mutations in the CFTR gene, providing the basis for understanding how CFTR works and how its various malfunctions cause disease. In News and Views Christopher Miller³ has summarized the evidence that CFTR is a Cl[−] channel, the context then being new evidence that some mutations cause decreased ion flow through CFTR⁴. The results of Tabcharani *et al.* are the latest triumph of a strategy that provides insights into CFTR function by analysing the effects of mild mutations.

For most non-biophysicists, the concept of multi-ion pores is probably esoteric. Ion channels can conduct millions of ions per second yet ions go through some regions of most channels in single file. Because of ion-ion repulsion in such a confined space, some pores can accommodate only a single ion at a time. But by design, some pores contain more than one

brane, the single-channel conductance is about 7 pS. Thiocyanate (SCN[−]) is more permeant than Cl[−], so that with equivalent concentrations of SCN[−] CFTR's conductance is around 10 pS. The anomaly appears when the two ions are mixed — replacement of 10 mM Cl[−] with the more permeant SCN[−] reduces the conductance from around 7 to about 2 pS.

Why does addition of a small amount of a more permeant anion actually cause the conductance to drop? At low concentrations, SCN[−] acts as a blocking ion. In multi-ion pores, powerful repulsive interactions occur among ions within the pore, enabling incoming ions to knock resident ions out of 'energy wells' or binding sites within the channel. In this way, flux through the pore is increased. Ions that bind tightly in the pore can be readily knocked out by other ions of the same type, but not by ions that bind less tightly. This kind of behaviour is exemplified by Ca²⁺ channels, which have refined this design principle to achieve high selectivity for Ca²⁺ ions without undue sacrifice of conductance (see Gary Yellen's News and Views article⁶). Thus sodium or lithium currents flow readily through Ca²⁺ channels in Ca²⁺-free solutions, but are blocked almost completely by the add-

ition of small amounts of Ca^{2+} . For CFTR, SCN^- causes a less dramatic block of Cl^- currents.

Where in CFTR does SCN^- bind? Nine positive residues located in or near the 12 predicted transmembrane helices of CFTR are conserved in all species examined to date, including shark, frog and human (see refs 7, 8 and figure). Remarkably, four of the nine have already been identified as sites for mutations causing cystic fibrosis (see ref. 4), and one of these, R347, has a location consistent with the electrical distance predicted from the voltage dependence of the block¹. In a demonstration that this positive charge is critical, Tabcharani *et al.* altered arginine to histidine (R347H), and controlled the charge by varying pH. At low pH, histidine is positively charged and R347H behaves like wild-type CFTR; but at higher pH, where the charge is neutralized, CFTR channels containing R347H had reduced conductance and the mole-fraction effect was abolished.

This finding is the latest contribution to knowledge of CFTR's channel function. We now know that CFTR forms a water-filled pore through the membrane⁹; that R347 is in the pore wall; that the positive charge of R347 confers multi-ion pore behaviour on CFTR; and that without such behaviour the conductance is reduced. Naturally occurring mutations of this site are found in patients with a milder form of cystic fibrosis. At least two other mild mutations, R117H and R334W, also reduce ion flow through CFTR⁴.

These experiments provide further evidence that reduced Cl^- conductance is the root cause of cystic fibrosis pathology. CFTR is now established as a Cl^- channel beyond reasonable doubt, and there is strong evidence that CFTR-mediated Cl^- conductance is essential for fluid secretion and salt absorption in the wet epithelia that are affected by the disease. When secretion falls below a critical minimum, blocked exocrine ducts can result. Through cascades of effects that are organ-specific and still poorly understood, entire structures like the vas deferens and pancreas can then be destroyed, and the lungs become susceptible to repeated infections.

This general explanation of cystic fibrosis is now widely accepted, but because CFTR resembles ATP-dependent trans-

porters such as P-glycoprotein, which may function as both a pump and a channel, an additional pump function for CFTR remains an intriguing possibility — and the loss of that function would presumably contribute to disease symptoms¹⁰. Most mutations cannot help in separating the channel and possible transport functions of CFTR, because they block the production or trafficking of CFTR protein^{11,12}, and mutations that interfere with ATP binding or hydrolysis would also affect any pumping function of CFTR. But missense

mutations that affect permeation of small anions would not necessarily affect the transport of other substances. Although it is conceivable that all such mutations alter other functions of CFTR, a simpler explanation is that reduced Cl^- conductance through CFTR is the primary physiological defect in cystic fibrosis. □

Jeffrey J. Wine is in the Cystic Fibrosis Research Laboratory, Jordan Hall, Stanford University, Stanford, California 94305, USA.

LANGUAGES

Mathematics in linguistics

Jared M. Diamond

AN exciting but controversial area in linguistics is the search for relationships between modern languages that diverged long ago from a common ancestral language¹⁻³. Such relationships would have major implications in anthropology, history and human genetics — for example, in understanding the origins of modern Europeans^{4,5}, American Indians² and of human language itself¹⁻³. But the search for such relationships is difficult, because the more distant the languages compared, the greater the risk that any resemblances will have arisen by chance, not by derivation from a common ancestor. These issues have been explored in a statistical null model by Ringe⁶, resulting in a response by Greenberg⁷ and a counter-response by Ringe⁸.

Linguists seek evidence of language relationships in grammatical similarities and cognates (cognates are words, in different languages, so similar in sound as well as meaning that they are presumed to have been derived from the same ancestral word). Sets of cognates are then examined for so-called recurrent sound correspondences: regular replacement of sound *X* in language A by sound *Y* in many words of language B. A familiar example is the correspondence of English *wh-* with German *w-*, as in the English/German synonyms *when/wann*, *white/weiss* and *whale/Wal*. These correspondences arose because modern English and German diverged from an ancestral Germanic language spoken around 2,000 years ago.

Linguists rarely use statistics to test the significance of postulated cognates or sound correspondences. That stems in part from the development of linguistics in the humanities rather than in the sciences. But also most linguists have concentrated on languages that are so similar that the evidence for a relationship between them is compelling even without statistical tests. Difficulties arise, however, in testing the postulated distant relationships

that are now attracting attention.

Words are composed of sequences of sound elements termed phonemes. If one compares consecutive phonemes of many synonyms from two languages, some recurrent sound correspondences are likely to occur simply by chance. Distantly related languages will in fact share few real sound correspondences. It is therefore tricky to evaluate whether the observed number of correspondences exceeds expectations based on chance. One expects the number of chance correspondences to increase with the number of synonyms compared, to decrease with the diversity of phonemes in the languages compared and to increase with inequality in relative frequencies of the phonemes.

Ringe⁶ begins by calculating the number of chance correspondences for word lists in arbitrary 'artificial languages' that he constructs, each with a specified set of phonemes at specified relative frequencies. He next shows that the number of observed correspondences exceeds those chance-based expectations for two pairs of languages whose relationships are well established, English/German and English/Latin. He finally applies the method to language pairs generally believed to be unrelated, such as English/Navajo.

As an example, consider Ringe's comparison of English and Turkish, which most linguists believe are unrelated but which belong to language families (Indo-European and Altaic respectively) that some linguists do consider are related¹⁻³. Comparing a list of 100 English words with their Turkish synonyms, Ringe finds 26 correspondences recurring two or more times between initial consonants, of which the most frequent is between English *b* and Turkish *k*. The six *b/k* synonyms are *bark/kabuk*, *belly/karn*, *bird/kuş*, *black/kara*, *blood/kan* and *bone/kemik*. Does this mean that the English *b* and Turkish *k* are derived from the same consonant in an ancestral language?

To answer this question, Ringe asks

1. Tabcharani, J. A. *et al.* *Nature* **366**, 79–82 (1993).
2. Riordan, J. A. *et al.* *Science* **245**, 1066–1073 (1989).
3. Miller, C. *Nature* **362**, 106 (1993).
4. Sheppard, D. A. *et al.* *Nature* **362**, 160–164 (1993).
5. Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, Massachusetts, 1992).
6. Yellen, G. *Nature* (in the press).
7. Tucker, S. J., Tannahill, D. & Higgins, C. F. *Hum. molec. Genet.* **1**, 77–82 (1992).
8. Cremonesi, L. *et al.* *Hum. Mut.* **1**, 314–319 (1992).
9. Hasegawa, H. *et al.* *Science* **258**, 1477–1479 (1992).
10. Gill, D. R. *et al.* *Cell* **71**, 23–32 (1992).
11. Cheng, S. H. *et al.* *Cell* **63**, 827–834 (1990).
12. Gregory, R. J. *et al.* *Molec. cell. Biol.* **11**, 3886–3893 (1991).

how often one would expect English and Turkish synonyms to share some initial consonant just by chance. English *b* and Turkish *k* happen to be the second-commonest initial consonants in their respective languages, so that one would expect them to correspond frequently in English/Turkish synonyms even if the two languages were totally unrelated. From actual relative frequencies of English and Turkish initial consonants, Ringe's null model predicts an average of 1.7 *b/k* matches by chance in a 100-word list. Ringe calculates the probability of obtaining six by chance as only around 0.01.

At first, that suggests that the *b/k* correspondence is significant. But there is the issue of 'sample space'⁹ to be taken into account. If 100 consonant pairs are examined, one would expect to find approximately one example of a chance correspondence that would individually be unlikely at a probability of only around 0.01. From this perspective, it is unsurprising that Ringe's matrix of 153 comparisons for English/Turkish initial consonants produces two correspondences that would appear significant in isolation; one of those happens to be *b/k*. Considerations of English etymologies also make a historical interpretation of the *b/k* correspondences unlikely: the English *b* words entered the language too late to have been derived from a common ancestor shared with Altaic languages. Naturally, this does not disprove the hypothesis that the Indo-European and Altaic language families are distantly related. Rather Ringe's calculations fail to uncover support for it.

Ringe next considers several methods that linguists employ to increase the number of historically significant cognates or sound correspondences detected. Among them are increasing the number of word pairs compared; increasing the number of potentially related languages compared simultaneously (so-called multilateral comparison); relaxing the requirement that the word pairs (or sets) compared must be identical synonyms; and relaxing the requirement of exact sound correspondences. Such techniques do indeed serve to increase the detected number of historically derived sound correspondences. But Ringe shows that, by the same token, they also increase the number of chance correspondences. The net result, in Ringe's model calculations, is that the task of detecting historical relationships becomes no easier and may even become harder. That conclusion poses a paradox, however. Such methods have been employed since the beginning of modern linguistics, and have yielded conclusions that are universally accepted because of overwhelming evidence (for example that modern Indo-European languages are related).

So does Ringe's model omit any techniques that help linguists distinguish

historically derived sound correspondences from chance ones? It does omit some, of course, because simpler models are necessary stepping-stones to more complex ones. A consideration discussed by both Ringe⁶ and Greenberg⁷ is that certain sound shifts with time (hence certain sound correspondences between pairs of languages derived from the same ancestral language) are observed much more often than others. For example, *p* often changes to *f*, and *t* often changes to *d*. When linguists see such correspondences between pairs of languages, they know by experience that those correspondences are likely to be historically derived, and that sound changes that have been rare in the history of human speech require stronger evidence before an apparent case can be taken seriously. Ringe^{6,8} notes that this fact is nevertheless difficult to formulate in a null model. It probably helps explain why linguists are more successful than a first-generation computer model in distinguishing significant sound correspondences from chance ones.

As a final example, Greenberg and Ruhlen¹ note that words similar to *milq* and meaning *swallow* or else *throat* appear in languages from at least six of the 11 proposed subfamilies of the Amerind language family, which in Greenberg's view encompasses most American Indian languages². The authors calculate the chances of an accidental match for these three consecutive consonants in six languages as only 1 in 10,000,000,000. This calculation, too, is a stepping-stone to a more complex and realistic calculation, because it omits considerations of sample space⁹ (such as number of words tested or number of languages examined per subfamily). Interestingly, similar words with related meanings appear in several Old World language families, including Indo-European, Afro-Asiatic, Uralic and Dravidian.

Whether a refined statistical treatment would strengthen or demolish this hint of very ancient linguistic relationships remains to be seen. But collaborations between statisticians and linguists might prove decisive in tackling that and similar issues. □

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90024, USA.

1. Greenberg, J. H. & Ruhlen, M. *Sci. Am.* **267** (5), 94 (1992).
2. Greenberg, J. H. *Language in the Americas* (Stanford University Press, 1987).
3. Ruhlen, M. *A Guide to the World's Languages*, Vol. 1 (Stanford University Press, 1987).
4. Cavalli-Sforza, L. L. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5620–5624 (1992).
5. Barbujani, G. & Sokal, R. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1816–1819 (1990).
6. Ringe, D. A. *Jr Trans. Am. Phil. Soc.* **82**, 1–110 (1992).
7. Greenberg, J. H. *Proc. Am. Phil. Soc.* **137**, 79–90 (1993).
8. Ringe, D. A. *Jr Proc. Am. Phil. Soc.* **137**, 91–109 (1993).
9. Campbell, L. et al. *Soc. Study Indigenous Lang. Newsl.* **12** (2), 7–8 (1993).

Internal medicine

"KEEP taking the tablets" says the doctor; but many of us lack the self-discipline to do so. A drug-implant, leaking the drug into the bloodstream all the time, avoids this problem. Daedalus is now taking this idea further.

He recalls that about a third of the human race are host to intestinal worms. Many of them don't even know it. Some tapeworms cause almost no symptoms, apart from reducing digestive efficiency. Tapeworm eggs have even been sold by unscrupulous vendors as slimming pills.

Even the simplest worms exude quite complicated molecules, such as the enzyme-inhibitors they use to prevent their host digesting them. (These also make it harder for the host to digest anything else: hence the slimming effect.) Many pharmaceuticals, of course, are simply enzyme-inhibitors. So, says Daedalus, it should be feasible to modify a worm to produce a specific pharmaceutical. Once installed in the gut, it would generate the drug automatically, steadily and indefinitely.

Many worms can be cultured *in vitro*, axenically. A DREADCO worm colony is now being mutated by irradiation, and genetically engineered more directly, to persuade the creatures to evolve or take up genes for the synthesis of various drugs. The end product should be a veritable stable of pharmaceutically accomplished intestinal worms.

Pharmacists will be delighted. One pill (a suitably coated worm egg) will now suffice for the most enduring symptoms. Worm therapy will be ideal for diseases such as schizophrenia and tuberculosis, whose long-term medication puts a heavy responsibility on the possibly unstable patient. And worm-borne contraception will free the most feckless teenager from the pregnancies that perpetuate her drab social position.

The technique might even be made quite automatic. Pinworms flourish and spread particularly well in insanitary, squalid, overcrowded urban conditions. Once introduced to such slums, contraceptive pinworms should spread widely, cutting back the local birthrate. When the population had fallen sufficiently, social regeneration would occur, sanitary standards would rise, the pinworms would be eliminated, and the inhabitants would regain their fertility.

This neat and simple piece of social engineering will, of course, bring down torrents of denunciation. Daedalus may be forced to substitute a strain of worm that has been sabotaged by deleting the digestion-inhibitors from its eggs. These will be digested by its host, preventing it from spreading through the community.

David Jones

Solar asteroid diversion

SIR — Ahrens and Harris¹ argue that the only practical means of diverting asteroids threatening to collide with the Earth is by the detonation of large neutron bombs near the asteroid (the 'stand-off' mode). But we have discovered an approach which may be better than the proposed nuclear technique. Our suggestion derives from the use of solar sails^{2,3} as a means of deflecting asteroids. Unfortunately, solar sails suffer both from very low thrust (roughly 10^{-5} N m⁻² of collector surface at 1 AU) and from the mechanical difficulty of tethering the sail to a rotating, perhaps precessing, asteroid. But they could be used to focus sunlight onto the surface of

equal to the applied thrust F divided by the mass m of the asteroid, $a_* = F/m$. For material evaporated by concentrated solar heat, the thrust is approximately $F = \alpha \dot{m} v_e$, where α is a numerical factor (taken here to be 0.5) that accounts for the hemispherical rather than linear expansion of the gas, $\dot{m}$ is the mass flow rate and v_e is the ejection velocity from the asteroid. The mass flow rate is essentially the rate of evaporation of surface material, given by the incident power divided by the heat of evaporation of the material, H_{vap} , plus whatever heat is absorbed in raising the temperature of the ejected gas. We take $H_{\text{vap}} = 15$ MJ kg⁻¹ for silicates and 3

MJ kg⁻¹ for ice, and assume an ejection velocity close to the sound speed, $v_e = 1$ km s⁻¹. This ejection velocity is also close to the average molecular speed of typical silicate vapour molecules (O₂, MgO, SiO) at the vaporization temperature⁷ of 1,500–2,000 °C (depending on pressure). In the case of water ice, another 1 MJ kg⁻¹ must be added to raise the average molecular speed to 1 km s⁻¹. For silicates, the resulting thrust is about 3.3×10^{-5} N W⁻¹ of collected energy, or 0.03 N m⁻² of solar collector at 1 AU — a factor of about 3,000 better than a pure solar sail, even for silicate asteroids (icy objects have a factor of four further advantage).

It is unlikely that the material boiled off an asteroid will be pure gas — the scale depth of thermal

penetration is approximately the same as expected particle sizes in the asteroid, so the material ejected will be a dusty gas. However, to the first order, this has no effect on the thrust: if the ejected material carries a mass fraction ϕ of dust, the ejection velocity is reduced by a factor $1-\phi$. On the other hand, the mass flow rate is increased by the factor $1/(1-\phi)$, so that their product, which equals the thrust, is unchanged.

Computations and experiments with high-power lasers in Moscow⁸ show that at power concentrations q of 10^4 kW m⁻² evaporation begins after 1 s. The delay is longer for lower concentrations: it goes roughly as q^{-2} . Nevertheless, at irradiances sufficient to bring the surface up to the evaporation temperatures for silicates, evaporation begins on a short timescale compared to the time for surface material to rotate through the illuminated spot on the asteroid's surface. The illuminated

spot thus does not need to track the surface: as the asteroid rotates under the spot evaporation continues steadily, the rotation serving to bring fresh material into the heated area.

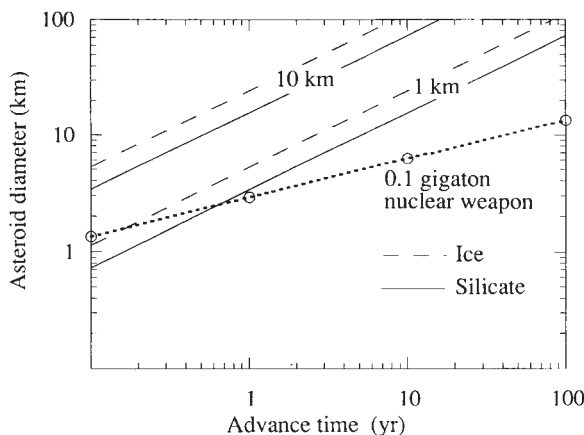
Bringing together these equations, we estimate that the diameter D of a circular solar collector capable of deflecting an asteroid of diameter d_a a time t before impact on Earth is given by:

$$D(\text{km}) = 0.16 \frac{d_a(\text{km})^{3/2}}{t(\text{yr})}$$

Thus, a 0.5-km diameter solar collector operating for a year can deflect asteroids up to 2.2 km in diameter. Estimating a mass of 5 g m⁻² for an aluminized mylar collector, which is well within current technology², a solar collector would weigh about 1 ton and could easily be lifted by the space shuttle. This same collector, functioning for a decade, could deflect a 10-km diameter asteroid. By contrast, the nuclear standoff proposal would require a 0.2–2.0 gigan nuclear weapon weighing more than 20 tons⁹, detonated 10 yr in advance¹.

The figure illustrates the deflection capability for 1- and 10-km diameter solar collectors, and for both silicate asteroids and icy comets. The plot also illustrates the deflection capability of a 0.1 gigan nuclear weapon. The time dependencies are different because nuclear deflection applies a single impulse, whereas solar evaporation affords a steady push. Long advance times thus favour the solar collector approach.

The advantages of the solar collector approach are that no tethers or special surface preparation on the asteroid are needed, the method provides a steady, gentle push that minimizes the danger of disrupting the asteroid, or even disrupting possible gravitationally-bound binary objects, and the system is reusable. Further, the solar collector is big, slow and fragile, so it could not easily be misused as a weapon. One byproduct of this technology could be an inexpensive method of interplanetary transport. Two mild disadvantages are that the solar collector must be delivered to the asteroid, although this may be accomplished by using the collector initially as a true solar sail to reach its target; and that a modest station-keeping propulsion system is required to keep the solar collector properly oriented and on target.



Asteroid deflection capabilities of solar collectors versus nuclear weapons. The plot shows the diameter of the asteroid (or comet) that can be deflected as a function of the time before impact. The pairs of solid and dashed lines are for silicate and icy bodies, respectively, that can be deflected by either 1- or 10-km diameter solar collectors. The heavy dotted curve with representative points is for the nuclear stand-off scenario employing a 0.1 gigan neutron bomb with an (optimistic) assumed conversion of 0.3 into neutron energy.

the asteroid to generate thrust as the asteroid's surface layers vaporize. In effect, the system is a solar-powered mass driver¹, but it does not require any equipment to be landed and operated on the asteroid's surface. This approach imitates the known effect of sunlight in generating the gas jets that perturb cometary orbits^{4–6}.

Ahrens and Harris¹ estimated that an asteroid on a collision course with the Earth would be deflected by a velocity impulse of 1 cm s⁻¹ applied 10 yr ahead of time. The velocity change for a steady thrust is slightly different, but to sufficient accuracy for small deflections it is given by the acceleration $a_* = 2 R_E / 3t^2$, where R_E is the Earth's radius and t is the time between the application of acceleration and impact, assuming that the thrust is applied along the direction of orbital motion (this accounts for the unfamiliar factor 1/3 in the equation for a_*). This acceleration is

1. Ahrens, T. J. & Harris, A. W. *Nature* **360**, 429–433 (1992).
2. Friedman, L. *Starsailing* (Wiley, New York, 1988).
3. Wright, J. L. *Space Sailing* (Gordon and Breach, Philadelphia, 1992).
4. Crifo, J. F. *Astr. Astrophys.* **187**, 438–450 (1987).
5. Peale, S. J. *Icarus* **82**, 36–49 (1989).
6. Whipple, F. L. *Astrophys. J.* **111**, 411–418 (1950).
7. Hashimoto, A. *Nature* **347**, 53–55 (1990).
8. Nemchinov, I. V., Medvedy, S. A. & Rybakov, V. A. *Hazards due to Comets and Asteroids* (ed. Gehrels, T.) (Conf. Univ. Arizona, Tucson) (in the press).
9. Canavan, G. H., Solem, J. C. & Rather, J. D. G. *Near-Earth-Object Interception Workshop* (Los Alamos National Lab., 1992).

The main problem for the solar collector is probably degradation of the collector surface by hot gas and dust ejected from the heated spot on the asteroid. This problem could be greatly ameliorated by the use of small, rugged secondary or tertiary mirrors near the asteroid, while the large primary collector could stand off some distance because of its long focal length (C. Steffens, personal communication).

H. J. Melosh

Lunar and Planetary Laboratory,
University of Arizona,
Tucson, Arizona 85721, USA

I. V. Nemchinov

Institute for Dynamics of Geospheres,
Russian Academy of Sciences,
Leninsky Prospekt 38, Bldg 6,
117334 Moscow, Russia

Viral burden in AIDS

SIR — The discussion¹⁻³ about the interpretation of the reports of a large viral burden in the lymph nodes of asymptomatic people infected with HIV is an argument about details which in essence are not disputed by either side. Essentially, HIV is able to kill or delete CD4 function by several direct and indirect mechanisms. The diversity of this virus and its cytopathic prospectives *in vitro* suggest that disease should rapidly ensue following infection. The fact that it usually does not and that many long-term survivors, as well as chimpanzees, do not seem to be succumbing to the effects of HIV infection, strongly suggests that CD4 tropism is not sufficient.

Sheppard *et al.*¹ are right in that immune activation is of paramount importance in determining disease outcome. In this regard, progression to AIDS has many features of disease of chronic activation, whether induced by viruses acting as superantigens (MMTV, MLV-Maids) or by chronic allogeneic stimuli⁴. One outcome of the generalized activation of the immune response is to increase the virus burden, therefore the mechanisms under debate are complementary. As several studies clearly show either a marked genetic predisposition or resistance to rapid HIV disease progression, and the virus load and pathogenicity can be separated in the SIV-infected macaque⁵ (a model for AIDS), the importance of the lymphoreticular viral burden cannot

merely be decided by simplistic mathematical models or by seeing how quickly CD4 cells can be killed in the test tube. It is therefore important to identify the host factors for progression and to define exactly which cells are infected in the lymph nodes and whether important cells, such as the follicular dendritic cells, need to be infected in order to activate or infect large numbers of lymphocytes, as has been suggested by Cameron *et al.*⁶.

The question of whether activation drives disease progression or whether it is merely a result of virus replication is so important that it is imperative to conduct selective immunosuppressive experiments in SIV-infected macaques. Prevention of SIV-induced disease by presuppression of the macaques before a live SIV challenge would compel us to view the issues under discussion in a new light.

A. G. Dalgleish

Department of Cellular and Molecular
Sciences,
St George's Hospital Medical School,
Cranmer Terrace,
London SW17 0RE, UK

SIR — It is becoming generally accepted that there is a larger viral load in HIV-infected individuals than hitherto thought. It is possible to calculate the total proviral load within an infected individual which, depending on the disease stage, is $>10^6$ – 10^{11} . But the proviral load provides a snapshot of events and gives little idea as to the dynamics of viral production and clearance. The viral load can be likened to a bucket standing under a tap: the limited size of the bucket is certainly important, but the consequences are ultimately determined by whether the tap was merely dripping or fully open. A wealth of data has described HIV infection in terms of chronic immune system activation: active polyclonal B-cell proliferation; intense HIV-specific cytotoxic T-lymphocyte effector frequencies; elevated peripheral CD8 cell counts; elevated serum β_2 microglobulin; and urine neopterin, to cite just a few.

Variation of the nucleotide sequence of HIV or SIV may yield pertinent temporal information as it represents a cumulative indicator of past events. HIV nucleotide sequence diversity increases over time from, in most cases, a homogeneous origin to reach an intrapatient diversity of $>10\%$. Longitudinal analyses of sequence diversity following experimental inoculation of macaques by SIV derived from molecularly cloned virus yielded nucleotide fixation rates of $\sim 1\%$ per site per year^{1,2}. Assuming a nucleotide substitution rate of ~ 0.1 – 1 per genome per cycle^{3,4}, approximately 500–5,000 consecutive rounds of replication would be necessary to achieve the observed 10% genetic variation between any two proviruses sampled at the same time. If the

viral burst size was merely two per cycle, a total of 2^{500} – $2^{5,000}$ (10^{150} – $10^{1,500}$) would result, well in excess of the number of hydrogen atoms in the Universe! Only if the mean burst size was restricted a hundredfold or more ($1.004^{5,000} \sim 10^9$, $1.04^{500} \sim 10^{8.5}$) could proviral loads approaching something like 10^6 – 10^{11} be achieved.

The most obvious explanation for the highly restricted viral burst would be an extraordinarily efficient clearance of virions and infected cells by the immune system. Most HIV-infected CD4⁺ T cells must therefore be cleared before virus can be produced, and loss of this cell type has always been the hallmark of HIV infection. The production of defective genomes could also limit amplification of HIV. But if the HIV mutation rate is <1 per genome per cycle, immune clearance is probably the dominant force in limiting the spread of HIV as opposed to the formation of defective viruses. Nonetheless, even though a defective provirus would be incapable of producing the full complement of viral proteins, as long as some were synthesized, the cell would be able to present HIV peptides in the context of MHC class I antigens, so attracting HIV-specific cytotoxic T cells — and CD4⁺ T-cell destruction.

Very little is known about the dynamics of the human immune system. Certainly the lifespan of memory (CD45RO) and naive (CD45RA) T cells is different⁵. Germinal centres come and go depending on the presence of antigen. T-cell homeostasis would seem possible, but the mechanics and implications for HIV disease are hard to assess⁶. Of course more is known about the mouse, where 30–40% of peripheral immunocompetent T and B cells are renewed every 3 days⁷. Much more information on the dynamics of the human immune system is needed.

Cumulative sequence variation reflects the repeated series of HIV/SIV CD4⁺ T-cell interactions and points to massive viral production and intense antiviral immune responses. Even though at any one time a minority of CD4⁺ T cells may be infected, given the numbers of replication cycles involved, and of course the time to accomplish these cycles, perhaps it is not too surprising that the immune system ultimately loses.

Simon Wain-Hobson

Unité de Rétrovirologie Moléculaire,
Institut Pasteur,
28 Rue du Dr Roux,
75724 Paris cedex 15, France

1. Sheppard, H. W., Ascher, M. S. & Krowka, J. F. *Nature* **364**, 291 (1993).
2. McLean, A. & Michie, C. *Nature* **365**, 301 (1993).
3. Garry, R. F. & Fermin, C. D. *Nature* **365**, 302 (1993).
4. Dalgleish, A. G. in *AIDS Research Reviews* (eds Koff, W. C., Wong-Staal, F. & Kennedy, R. C.) 73–93 (Dekker, New York, 1993).
5. Dalgleish, A. G. *Curr. Opin. Immun.* **5**(4) 608–614 (1993).
6. Cameron, P. U. *et al. Clin. exp. Immun.* **88**, 226–236 (1992).

1. Burns, D. P. W. & Desrosiers, R. C. *J. Virol.* **65**, 1843–1854 (1991).
2. Johnson, P. R. *et al. Virology* **185**, 217–228 (1991).
3. Pathak, V. K. & Temin, H. M. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6019–6023 (1990).
4. Leider, J. M., Palese, P. & Smith, F. I. *J. Virol.* **62**, 3084–3091 (1988).
5. Michie, C. *et al. Nature* **360**, 264–265 (1992).
6. Margolick, J. B. *et al. J. AIDS* **6**, 153–161 (1993).
7. Freitas, A. A. & Rocha, B. B. *Immun. Today* **14**, 25–29 (1993).

Skin cancer and UV radiation

SIR — Ultraviolet-B radiation (280–320 nm) has been implicated in induction of skin cancers in fair-skinned people. Reductions in stratospheric ozone (O_3) are expected to allow more solar ultraviolet-B to reach the Earth's surface. Calculated increases corresponding to O_3 reductions over 1979–89 have already been given^{1,2}; we have now updated those estimates using O_3 measured until the end of December 1992 by the Total Ozone Mapping Spectrometer (TOMS) aboard the Nimbus 7 satellite³. We also combine the ultraviolet-B increases with new data on skin cancer induction by ultraviolet radiation to estimate the corresponding increase in skin-cancer incidence.

We used a radiative transfer model to calculate the increase in ultraviolet-B, assuming that the overhead ozone column³ is the only atmospheric parameter to change¹. Biologically effective ultraviolet doses are computed as

$$D = \int F(\lambda, t) B(\lambda) d\lambda dt$$

where D is the dose received over time t , $F(\lambda, t)$ is the solar spectral irradiance at the surface and $B(\lambda)$ is the relative spectral sensitivity (action spectrum) of a specific biological effect. We computed doses using monthly average TOMS data (Jan 1979 to Dec 1992) from pole to pole in 10° latitude bands, then integrated over each year. We considered three action spectra relevant to the induction of skin cancer: for ultraviolet-induced skin cancer in mice⁴, for erythema induction in fair-skinned people⁵ and for DNA damage⁶. Obviously, no direct action spectrum is available for carcinogenesis in human skin.

Increases in annually integrated ultraviolet doses, shown in the table, are within one standard deviation (1σ) of zero only within about 10° latitude of the Equator, and become significant at the 2σ level polewards of 35° in both hemispheres. Relative to the 1979–89 changes¹, mid- and high-latitude trends are generally persisting and now carry substantially smaller uncertainties. A noteworthy difference is the appearance of trends which differ from zero by 1σ or more at latitudes within 15° of the Equator.

The relation between skin cancer incidence and ultraviolet dose is generally nonlinear, and is commonly represented by an exponential model⁷

$$\text{incidence} = \text{constant} \times (D)^{\text{BAF}}$$

where BAF is the so-called biological amplification factor, and is a function of, among other things, the cancer type and the action spectrum used to compute the ultraviolet dose. In a recent review⁷ of

Latitude	Total ozone	Erythema induction dose	DNA damage dose	Skin cancer dose	Basal cell carcinoma incidence	Squamous cell carcinoma incidence
85 N	-8.8±3.2	7.1±1.7	14.8±3.6	10.6±2.5	15.1±5.6	28.5±11.2
75 N	-9.0±2.9	7.6±1.7	14.9±3.3	10.8±2.4	15.4±5.8	29.1±11.5
65 N	-7.4±1.7	7.6±1.7	14.1±3.2	10.3±2.3	14.7±5.6	27.7±11.0
55 N	-7.4±1.3	7.2±1.7	12.9±3.2	9.5±2.3	13.5±5.3	25.4±10.3
45 N	-6.6±1.2	6.5±1.7	10.9±3.0	8.1±2.2	11.6±4.7	21.6±9.0
35 N	-4.8±1.4	5.0±1.8	8.2±3.0	6.1±2.2	8.6±4.0	16.0±7.6
25 N	-2.7±1.5	3.0±1.9	4.8±3.1	3.5±2.2	5.0±3.5	9.0±6.4
15 N	-1.5±1.1	1.7±1.4	2.6±2.3	1.9±1.6	2.7±2.4	4.8±4.4
5 N	-0.6±1.6	0.7±1.9	1.2±3.0	0.8±2.2	1.2±3.1	2.1±5.5
5 S	-1.1±1.4	1.2±1.7	2.0±3.0	1.4±2.0	2.0±2.8	3.6±5.2
15 S	-1.9±1.3	2.2±1.5	3.5±2.4	2.5±1.7	3.6±2.6	6.5±4.8
25 S	-2.6±1.6	3.1±1.7	5.0±2.4	3.6±1.9	5.1±3.1	9.2±5.8
35 S	-4.0±1.6	4.8±1.6	7.9±2.6	5.7±1.9	8.1±3.6	14.9±6.8
45 S	-5.6±1.4	7.2±1.6	12.5±2.7	8.9±1.9	12.7±4.8	23.9±9.2
55 S	-9.0±1.5	10.9±2.0	19.7±3.6	14.2±2.6	20.4±7.4	39.3±15.1
65 S	-15.0±2.0	16.3±3.0	30.5±5.8	21.9±4.1	31.9±12.2	64.0±26.6
75 S	-19.5±2.6	24.1±5.4	49.8±12.0	34.0±7.7	50.6±21.4	107.7±52.0
85 S	-21.1±3.0	31.0±6.8	72.0±17.6	46.5±10.5	70.6±31.2	159.6±83.6

Increases in doses are evaluated over the 14-year data record, expressed as per cent relative to the 1979 intercept. Uncertainties are 1σ . See text for description of action spectra used in the dose calculations.

The figures for carcinoma incidence are based on skin-cancer dose increase and biological amplification factors given in text.

epidemiological determinations of the BAF for fair-skinned people, values of 1.4 ± 0.4 were given for basal cell carcinomas and 2.5 ± 0.7 for squamous cell carcinomas (for cutaneous malignant melanoma, the scatter in the BAF determinations is still too large to allow for a reliable estimate of the increase of incidence). Application of these factors to the O_3 -induced-skin-cancer dose increases leads to estimation of increases in the incidence of the non-melanoma skin cancers at middle and high latitudes of both hemispheres, as shown in the last two columns of the table. Note that the largest per cent increases are at high latitudes where baseline incidence of skin cancer is usually small⁷.

Our calculation assumes that behavioural and demographic risk factors remain constant; that the 1979–92 increases in ultraviolet are sustained without further increase over the several decades required for skin-cancer development; and that the BAF model is independent of latitude and extent of ozone depletion. Some of these assumptions may be quite conservative, particularly the assumption that no additional ultraviolet increases will occur over the next few decades, while the assumption of no behavioural change may be too pessimistic. With stratospheric chlorine expected to increase until the turn of the century and not return to 1990 levels for several decades, it seems likely that ultraviolet-B will remain above

natural levels, and perhaps increase further, well into the next century.

An important additional note is that the various ultraviolet-B monitoring networks now being planned will begin data collection in an environment which is already significantly perturbed from its natural baseline level at middle and high latitudes.

Sasha Madronich

National Center for Atmospheric Research,
Boulder, Colorado 80307, USA

Frank R. de Gruijl

Department of Dermatology,
University of Utrecht,
Utrecht 3508 GA,
The Netherlands

Fitness-related dispersal

SIR — Brown¹ offered an alternative explanation to my conclusion² that lifetime reproductive success payoffs, related to territory quality, predicted observed dispersal behaviour in the group-territorial bird the Seychelles warbler *Acrocephalus sechellensis*. He suggested that breeding vacancies on low-, medium- and high-quality territories were mainly filled by nonbreeders from territories of the same quality because of the easier access to these vacancies (for example high-quality territories almost never border on low-quality habitat and because there are fewer medium-quality territories — the 'area-restricted' hypothesis). Another suggestion, which in my view my data support, is that the dispersal to high-quality territories could be due to higher competition because there are more nonbreeders present in these territories (the 'queue' hypothesis).

During the removal experiments, as I

1. Madronich, S. *Geophys. Res. Lett.* **19**, 37–40 (1992).
2. Crutzen, P. J. *Nature* **356**, 104–105 (1992).
3. Gleason, J. et al. *Science* **260**, 523–526 (1993).
4. de Gruijl, F. R. et al. *Cancer Res.* **53**, 53–60 (1993).
5. McKinlay, A. F. & Diffey, B. L. in *Human Exposure to Ultraviolet Radiation: Risk and Regulations* (eds Passchier, W. F. & Bosnjakovic, B. F. M.) 83–87 (Excerpta Medica, Int. Congr. Ser. 744, Amsterdam, 1987).
6. Setlow, R. B. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3363–3366 (1974).
7. Krieger, A., Armstrong, B. K., Jones, M. E. & Burton, R. C. *Health. Solar UV Radiation and Environmental Change* (Int. Agency Res. Cancer, WHO, Lyon, 1993).

mentioned in my paper², territory quality-related dispersal was not biased by a skew in the distribution of surrounding alloparents on different quality territories. Ten low-quality vacancies were adjacent to territories of medium and high quality with one or more prebreeding birds available to fill the vacancy, but these were filled only by young born on low-quality territories, sometimes from the other side of the island. Five vacancies were filled by males, who had been floaters for 2½ (2 birds), 3, 3½ and 6 years, respectively.

The larger number of nonbreeders in adjacent high-quality territories did not disperse to low-quality habitats; they all stayed in their own territory. Similarly, all 14 high-quality vacancies were adjacent to lower-quality territories with prebreeders which could potentially have filled the vacancy, but, as I stated² these birds were not able to fill these vacancies due to high competition as the result of the larger number of nonbreeders present in high-quality territories. Competition for breeding vacancies, measured as the inverse of time necessary to fill a vacancy, significantly increased with quality³. On average, vacancies in high-quality territories were filled within 16 hours, 1.9 and 4.3 times as fast as vacancies in medium- and low-quality territories, respectively (the 'queue' hypothesis).

The 'area-restricted' hypothesis can also be rejected by comparing the dispersal distances of individuals filling experimental vacancies in territories of the same quality as their natal territories (including only those vacancies in territories which border on territories of the same quality with nonbreeders present to fill the vacancies): 30.8% of breeders in high-quality territories were from adjacent territories ($n = 13$); for medium-quality territory and low-quality territory this figure was 28.6% ($n = 7$) and 30.8% ($n = 13$), respectively ($\chi^2 = 0.013$, d.f. = 2, not significant).

In conclusion, variation in territory quality (partially explained by the 'queue' hypothesis) is involved in the evolution of cooperative breeding in the Seychelles warbler. The results cannot be explained by the 'area-restricted' hypothesis.

Jan Komdeur

National Environmental Research Institute,
Department of Wildlife Ecology,
Kalø, Grenåvej 12,
DK-8410 Rønde, Denmark

1. Brown, J. L. *Nature* **363**, 308–309 (1993).
2. Komdeur, J. *Nature* **358**, 493–495 (1992).
3. Komdeur, J. thesis (Univ. Cambridge, 1991).

■ Several readers have pointed out that the K number for Mozart's sonata for two pianos in D major is 448 and not 488, as published in the Scientific Correspondence by Rauscher *et al.* (*Nature* **365**, 611; 1993). □

Dinosaur blood cells rediscovered

SIR — Although the recent find of possible blood cells in a 65 million-year-old *Tyrannosaurus* has had wide publicity^{1,2}, readers may not be aware that similar bodies were first reported³ 86 years ago, from a Bernissart *Iguanodon* roughly twice as old, or that possible osteoclasts⁴ were noted 70 years ago in a still older Upper Jurassic (Morrison) sauropod (? *Apatosaurus*).

Seitz's "*Vergleichende Studien*"³ of 1907 is one of the two early classics of palaeontological bone histology, giving detailed descriptions of bone tissues from 40 fossil reptiles, including 12 dinosaurs. In *I. bernissartensis*, Seitz found that some Haversian (vascular) canals contained small rounded "double contoured" bodies, occurring scattered individually along the canals or locally forming compact masses. On the assurance of the botanist Solereder that these were not spores, Seitz surmised that they were blood corpuscles, whose filling the canals locally could represent coagulation or clotting. Supporting this, accumulations of small reddish crystals could represent haematoidin (a breakdown product of haemoglobin). Seitz, however, offered these observations "with reservation".

Moodie⁴ then found similar possible corpuscles in a sauropod, but suggested that they could instead be osteoclasts, because trabeculae between vascular spaces showed Howship's lacunae which are formed when these cells resorb bone. This notion was later adopted by Swinton⁵, who held that there were no certain records of blood cells from dinosaurs, and that at least some supposed examples were probably iron-stained osteoclasts. Both authors^{4,5} also doubted that the crystals seen by Seitz could represent haematoidin.

So, did Seitz find blood cells in *Iguanodon*? It seems reasonably likely that he did, despite Swinton's opinion. Osteoclasts are large multinucleate cells, seen chiefly within resorption spaces and other regions of active reconstruction, and not within vascular canals that are not undergoing enlargement. Moodie's reference to trabeculae implies that his examples were seen in cancellous bone, which is usually reconstructed bone in dinosaurs except near epiphyses⁶; but Seitz's were found instead in vascular canals in compact bone, and not in resorption spaces. His figure shows no Howship's lacunae, and he reported none from other parts, although introductory text shows that he knew their significance. One would also not expect to find densely packed osteoclasts filling vascular canals in the manner of coagulated blood cells, as seen, for

example, in post-mortem sections. Thus, Seitz's view fits his evidence better than Swinton's.

Do such objects have any particular value in the search for dinosaurian DNA? Cells of either type described above could yield DNA fragments, as sauropsid erythrocytes are nucleate, and could provide a readier source than bone. Any bone retaining organic traces is a potential source, but recovered organic content will always be derived mainly from the original collagen and mucopolysaccharide. But if actual cells could be isolated, even in a mineralized condition, this problem would be largely avoided. A different and more common type of object might, however, be more useful, as potentially more resistant to leaching or contamination. In some bones, osteocyte lacunae and canaliculi have a thick enough inert mucopolysaccharide lining for maceration residues to yield osteocyte pseudomorphs, which are known from bones as old as Carboniferous⁷. Such bodies might prove to yield 'packeted' DNA traces, insulated from changes to which soft-walled cells would be susceptible; and they could be obtainable from *Tyrannosaurus*, as Pawlicki⁸ found mucopolysaccharide concentrated around lacunae in the closely allied *Tarbosaurus*.

Does the recovery of dinosaurian DNA have any scientific value? Just possibly, it might reveal whether dinosaurian extinction could have been due to climatic change upsetting sex ratios⁹, as this would have happened only if sex was determined by incubation temperatures, and not chromosomally. But we would still not know what actually caused dinosaur extinction.

R. E. H. Reid

School of Geosciences,
The Queen's University of Belfast,
Belfast BT7 1NN, UK

1. Morell, V. *Science* **261**, 160–162 (1993).
2. Kiernan, V. *New Scientist* 1881, 5 (1993).
3. Seitz, A. L. L. *Nova Acta Acad. Caesar. Leop. Carol.* **87**, 230–370 (1907).
4. Moodie, R. L. *Palaeopathology* (Univ. Illinois Press, Urbana, 1923).
5. Swinton, W. E. *The Dinosaurs* (Murby, London, 1934).
6. Reid, R. E. H. in *Zool. Soc. Lond. Symp.* (ed. Ferguson, M. W. J.) **52**, 629–663 (1984).
7. Halstead, L. B. *Vertebrate Hard Tissues* (Wykeham Science Ser., London, 1974).
8. Pawlicki, R. *Acta histochem.* **58**, 75–78 (1977).
9. Ferguson, M. W. J. & Joanen, T. *Nature* **296**, 850–852 (1982).

Scientific Correspondence

Scientific Correspondence is a relatively informal section of *Nature* in which matters of general scientific interest, not necessarily those arising from papers appearing in *Nature*, are published. Because there is space to print only a small proportion of the letters received, priority is usually given according to general interest and topicality, to contributions of fewer than 500 words, and to contributions using simple language.

Late night thoughts

Jacqueline Reynolds

The Biological Century: Friday Evening Talks at the Marine Biological Laboratory. Edited by R. Barlow, J. Dowling and G. Weissmann. *Harvard University Press: 1993.* Pp. 289. \$45, £35.95.

As the nineteenth century came to an end, the United States was emerging as a centre for scientific research in areas traditionally dominated by Western Europe. The beginnings of an American school of anthropology came from Lewis Henry Morgan and Franz Boas; physics and mathematics had the shining star of J. Willard Gibbs; T. H. Morgan and E. B. Wilson brought new life to genetics and development. The 'colonials' no longer simply absorbed the scientific wisdom of Britain and the Continent, but now transmitted their own new discoveries and ideas across the Atlantic. In the midst of this intellectual activity, members of the Boston Society of Natural History and the Woman's Education Association of Boston established in 1880 what is now the Marine Biological Laboratory at Woods Hole, Massachusetts — originally to "teach biology to teachers", but soon to become a popular research station for biologists of all persuasions.

To celebrate the century since C. O. Whitman was appointed its first director, the laboratory has collected a series of its traditional "Friday Night Lectures" given in honour of some of the biological scientists whose studies are the foundation of much of the research carried out in Woods Hole today. Nine of the eleven essays are summaries of the advances in a specific biological field since the inception of the laboratory, and each is preceded by brief biographies of the earlier scientist(s) and the lecturer.

The subjects covered are wide ranging and represent a multitude of disciplines: for example, genetic mapping from T. H. Morgan and A. H. Sturtevant to the Human Genome Project; generation of membrane potentials from W. J. Osterhout to the amino-acid sequence of channels and ion pumps; marine ecology from S. F. Baird's questions in 1871 about the demise of some commercial fish to investigations in the 1980s of the effect of nutrient supplies on algal growth. Some of these overviews give a real feel for earlier biological research and its influence on modern science, but occasionally the excitement of today's theories overwhelm the lecturer and are presented with as much surety as fundamental laws that have stood the tests of time.

There are entertaining and thought-provoking titbits throughout. For example, we learn that E. G. Conklin had a life before he delved into developmental bi-

ology. He was hired by Rusk University between 1885 and 1888 to teach Latin, Greek, English, elocution, history and all of the sciences. Is there a polymath today who could meet such a challenge? In a section on neural transmission, Torsten Wiesel compares Keffer Hartline, the pioneer of retinal physiology, to a novelist writing about a single city at progressively deeper levels. By contrast, Stephen Kuf-



Baywatch — Marine Biological Laboratory students on an invertebrate-zoology collecting trip at Quissett harbour, near Woods Hole, in 1897. Gertrude Stein is in the centre foreground.

fler, renowned for his work in neural transmission, is likened to another novelist pursuing the same theme in many different settings. "Science, like art, is a medium through which a spectrum of individual styles may find expression." How true, but how little appreciated by the 'big business' science of today. Elsewhere, Joshua Lederberg, relating the history of genetic mapping, concludes with a thoughtful discussion of the huge cost and perceived benefits of the Human Genome Project. He reminds us that "to have merely sequenced the DNA is just the tip of the iceberg". The enormous effort required to understand the functional aspects of each gene and its product should not be underestimated, and it is only with this understanding that real advances will come.

The best parts of this volume are two purely historical essays that should appeal to scientists and laymen alike. Gerald Weissmann writes with enthusiasm about the unlikely connection between Jacques Loeb, the founder of the mechanistic school of American biology, and Gertrude Stein who, as an aspiring medical student, spent the summer of 1897 at

Woods Hole before going off to Johns Hopkins Medical School. Did Loeb's reductionist philosophy influence the young Stein in her later revolutionary use of words? Were her experiences at the laboratory related to (or responsible for) the abandonment of a medical career for the literary scene? Weissmann makes a persuasive case for these possibilities: "As Loeb reduced the structure of living things to the bare molecular world, so Stein reduced language to its bare molecular level where phonemes throb to their own rhythm".

The second historical account is a portrait (warts and all) of Albert Szent-Györgi, written by his friend and associate, Benjamin Kaminer. Szent-Györgi first went to the laboratory in 1947 and found a congenial home there for his

sometimes eccentric approach to research. His "flights of imagination and hasty pronouncements" were not always popular with other scientists, but his sense of humour and joy of life endeared him to many. Socially as well as intellectually, he made an impact on the Woods Hole community, taking up water skiing at the age of 80 and leading the Fourth of July parade through the village's main street.

Many changes have occurred since the founding of the Marine Biological Laboratory, in both the level of our understanding of biological systems and the manner in which science is conducted. A more comprehensive historical study of this unique institution and those associated with it over the years would be welcomed by scientists and historians alike — a portrait in miniature of the development of science in an environment quite different from the grand, well-endowed universities. For now, we will have to be satisfied with the glimpses afforded in this collection of lectures. □

Jacqueline Reynolds is at Tarlswold, Back Lane, Easingwold, North Yorkshire YO6 3B6, UK.

Gravitational breakdown

R. Cowsik

The Rise and Fall of the Fifth Force: Discovery, Pursuit, and Justification in Modern Physics. By Allan Franklin. *American Institute of Physics*: 1993. Pp. 141. \$29.95, £21. (Distributed in the United Kingdom by Oxford University Press.)

How do scientists come to formulate new scientific hypotheses and in what context? How do these hypotheses get rejected or, in some rare instances, canonized as new laws of nature? What is the driving force behind scientific endeavour? Allan Franklin attempts to answer these questions by examining the suggestion made by Ephraim Fischbach and colleagues in 1986 of a possible extra component to gravity.

Whereas conventional gravity generates the same acceleration on all bodies, irrespective of their composition, this new force would distinguish between different substances and cause them to move at different speeds. Moreover, unlike gravitation, which is a long-range force whose strength decreases in proportion to the inverse square of distance, the new force was proposed to decrease exponentially with distance (similar to the decrease in nuclear forces as suggested by Yukawa, albeit with a typical range of 10–100 metres).

Fischbach and colleagues based their suggestion on several apparently disjointed clues: in 1955, Lee and Yang had discussed the possibility of a long-range force that couples baryons akin to the electrostatic coulomb force that couples electric charges. Some ten years later, Lee, who with Yang had pioneered the discovery that parity was not a universal symmetry in nature but was maximally violated in weak interactions, attempted to save the combined symmetry of charge conjugation (particle-to-antiparticle interchange) and parity in the decay of the K_L^0 meson. Violation of this combined symmetry — which is equivalent to the violation of time-reversal invariance — was, according to Lee and others, only seemingly so; the real cause of the apparent CP-violation was a long-range force that couples hypercharge. Other clues came from experiments on gravitation, notably those of Eötvös; his results had so impressed Einstein that they led him to formulate the principle of equivalence on which, indeed, all metric theories of gravitation are based.

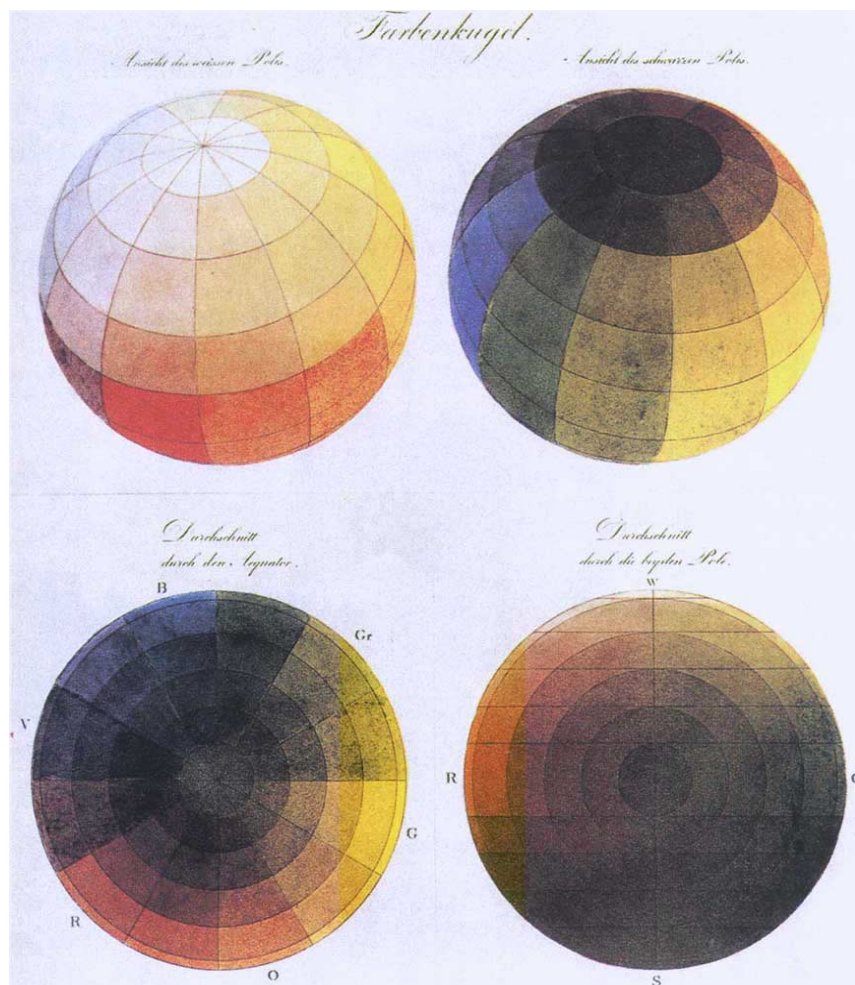
Yet despite these exciting beginnings, and probably because of them, within about five years many new experiments were done that led to the firm conclusion

that such a new force was not there in nature, at least not at strengths greater than one part in ten thousand of gravity.

This brief interlude in the contemporary history of science is well documented by Franklin, and should be of much interest to those concerned with the philosophy of science and epistemology. Classical epistemology based on complementary and often confrontational ideas of empiricism and rationalism similarly underwent a huge change in the post-newtonian period and indeed had to be modified substantially to accommodate Einstein's idea of relativity, spacetime and gravitation. With such an evocative title, however, one expects to find a book of

more majestic dimensions with a detailed analysis of the episode in terms of the competing ideas in the field of philosophy of science and epistemology. Instead, the author provides a detailed review not merely of the scientific papers and conference discussions, but also of the informal electronic-mail exchanges of the original collaborating workers. This meticulous chronicle, with its comments on the basic questions of scientific methodology and verification of scientific truth, should therefore constitute an important reference for more extensive studies. □

R. Cowsik is in the Indian Institute of Astrophysics, Bangalore 560034, India.



NEWTON'S *Opticks* sought to put the study of light and colour on an objective, quantitative basis. But it was increasingly the subjective colour phenomena, such as colour contrasts and the Purkinje shift, that occupied the attention of scientists who developed the modern study of colour vision. In particular, Goethe's *Theory of Colours* (1810), a polemic against Newton, made an important contribution to the developing science of the physiology and psychology of perception in the nineteenth and twentieth centuries. The poet's influence also extended to the Romantic artists who sought to extract new meanings for colours from their positions in space. Runge's *Colour Sphere* (1810), shown here, was one of the earliest attempts by a painter to coordinate hues and values (light–dark content) into a coherent whole. This picture appears in John Gage's *Colour and Culture*, an immensely detailed and sweeping analysis of colour in Western culture from the ancient Greeks to the present. Copiously illustrated, it will prove irresistible to artists and scientists alike. Thames and Hudson, £38.

Barriers to gene flow

R. K. Butlin

Species Evolution: The Role of Chromosome Change. By Max King. Cambridge University Press: 1993. Pp. 336. £40, \$59.95.

FIFTEEN years ago Michael White published *Modes of Speciation*, in which he argued that chromosomal change plays an important part in the origin of species. Max King believes that the book, a masterful synthesis of data in support of this position, has received far less attention than it deserves and that the role of chromosomal change has been played down consistently by the "essentially North American coterie of like-minded geneticists" who, from his Australian perspective, currently dominate the field. He sets out on a crusade to reinstate cytogenetics (or cytogeneticists?) in its rightful place and to show what progress has been made since 1978 despite the perceived neglect.

All models of speciation need to overcome the same central problem: a mutant allele that contributes to reproductive isolation must be at a selective disadvantage when it first appears in a population. Non-chromosomal models have various ways around this obstacle, but chromosomal theories generally tackle it head on. The case for a primary role of chromosomes in speciation is based on the strong fitness reduction in heterozygotes produced by some types of rearrangement that generate, in one step, a substantial barrier to gene flow. King offers a familiar pair of circumstances to overcome the problem of the spread of such a rearrangement: inbreeding in small populations and meiotic drive. It is not clear why these conditions should apply to chromosomal rearrangements and not to single-locus mutations with large effects on the fitness of heterozygotes. Although King argues that the chromosomal mutation rate is high (greater than 1 in 500), this is not high relative to the per-genome rate for single-locus mutations. His claim that meiotic drive is an important evolutionary force may be much less controversial now than in 1978, but the link between meiotic drive and underdominant chromosomal rearrangements remains tenuous.

Few would doubt the evidence, which King presents at length, that some chromosomal rearrangements contribute to selection against interspecific hybrids. The question is whether the chromosomal change is the primary event in the speciation process. Here, King relies on evidence from allozymes to measure

'genic' divergence and marshals examples in which chromosomal changes characterize taxa that lack 'genic' differentiation. The example of the alpine grasshopper, *Podisma pedestris*, clearly shows the danger of this argument. There is strong evidence that many genes contribute to postzygotic isolation between races despite the absence of allozymic differentiation, but the overall isolation is much greater than the isolation produced by the chromosomal fusion that characterizes them. It seems likely that the main difference between chromosomal and genic contributions to reproductive isolation is simply that chromosomal changes can be seen under a microscope: either can evolve without allozymic divergence.

In his final comments, King laments "a failure to integrate specific findings into a common evolutionary perspective". Sadly, this is precisely the problem with his book: he makes a case, and makes it well, for the existence of chromosomal speciation but he does not provide arguments for its importance relative to other modes of speciation. Like its predecessor, the book deserves a readership among workers in the field, but for a general coverage of 'species evolution' one will have to look elsewhere. □

R. K. Butlin is in the Department of Genetics, University of Leeds, Leeds LS2 9JT, UK.

Star wars

Joel A. Gwinn

Pauper and Prince: Ritchey, Hale, and Big American Telescopes. By Donald E. Osterbrock. University of Arizona Press: 1993. Pp. 359. \$45.

THE giant telescopes constructed during the first half of the twentieth century have revealed much about our place in the cosmos. We now know a great deal about the astrophysics of the Sun and its system of planets and the relative position of our Solar System in the Milky Way; and spectral classification of stars has led to new methods of determining stellar distances and, through establishment of intergalactic distances, to the modern cosmological view of a Universe expanding from an initial Big Bang.

The 1888 dedication of Lick Observatory, with its 36-inch refracting telescope, marked the beginning of a trend of building huge telescopes for astronomical research. The wealthy individuals and foundations able to provide the requisite funds were typically motivated by a desire to build "the world's largest telescope". In the astronomical community, George

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Hale: nobility of wealth.

Ellery Hale is well known for his success in identifying potential donors and persuading them to support the next venture.

Hale was also an accomplished research astronomer in his own right. While Hale was still a student at the Massachusetts Institute of Technology, his father provided for him, on a lot next to the family home, his own fully equipped observatory, the Kenwood Physical Laboratory. Over the years, this wealthy Chicago businessman provided telescopes and observatories, and hired assistants to further his son's career in astronomy. American nobility is a nobility of wealth, and George Ellery Hale was a true member of this nobility, with all the status and privileges that come with it, together with the neurosis and arrogance, as we find out in this remarkable book.

As a newly appointed associate professor of astrophysics at the University of Chicago, Hale obtained two 40-inch glass disks, and, together with W. R. Harper, the president of the university, persuaded Charles T. Yerkes, the Chicago street-car magnate, to pay for a new observatory based on a new refracting telescope even larger than the 36-inch giant at Lick Observatory.

Born almost five years before Hale, George Willis Ritchey was a descendant of Scotch-Irish immigrant craftsmen. He was not entitled to the rights and privileges of American nobility of wealth. All his life, he lacked the money needed to pursue his creative visions, prohibitively expensive owing to his standards of perfection.

Working as an optician on a salary provided by Hale's father, Ritchey developed a symbiotic relationship with Hale, first at Kenwood, and then at Yerkes, where he obtained better photo-

graphs of the Moon than any previously made. At Yerkes he built the first professional 24-inch reflector and began work on a 60-inch reflector. Hale's father paid for and retained ownership of the 60-inch disks. This enabled Hale to take them along when he and Ritchey broke with the University of Chicago to form the new Mount Wilson Observatory, supported by the Carnegie Foundation and directed by Hale from 1904 to 1923.

Ritchey set new standards of excellence in astrophotography, developing and refining his methods for grinding, shaping and testing large mirrors. When the 60-inch telescope was completed in 1908, Hale pronounced it "practically perfect". Ritchey and others made many important observations with the 60-inch reflector, obtaining many revealing photographs of "nebulae" that we now know to be galaxies outside the Milky Way. In 1917 he found a nova or "new star" in one of these spiral galaxies, showing it to be a star system distant from our own. As the excitement of these new discoveries spread, Ritchey became recognized as the leading authority on large reflecting telescopes.

Even before the 60-inch was completed, Hale and Ritchey had begun to create a 100-inch reflector and had persuaded John D. Hooker to pay for it. The next-to-impossible task of obtaining a suitable 100-inch mirror blank fell largely to Ritchey. The mutually beneficial co-operative relationship between Ritchey and Hale had already begun to deteriorate, however. This eventually led to Ritchey's dismissal from the Mount Wilson Observatory, and to many unhappy years of ostracism from the astronomical community.

New in paperback

Did Darwin Get It Right? Essays on Games, Sex and Evolution by John Maynard Smith. Penguin, £6.99. See *Nature* **339**, 438 (1989).

The Theory of Evolution by John Maynard Smith. Now with a forward by Richard Dawkins. Canto/CUP, £7.95, \$11.95.

Artificial Life: The Quest for a New Creation by Steven Levy. Penguin, £8.99. See *Nature* **360**, 25 (1992).

Ecological Imperialism by Alfred W. Crosby. CUP, £7.95, \$11.95. See *Nature* **323**, 587 (1986).

Artificial Worlds: A Journey in Hype and Hyperreality by Benjamin Woolley. Penguin, £5.99. See *Nature* **362**, 421 (1993).

The Unnatural Nature of Science by Lewis Wolpert. Faber, £6.99. See *Nature* **360**, 378 (1992).

The Transformed Cell: Unlocking the Mysteries of Cancer by Steven A. Rosenberg and John M. Barry. Avon, \$12.50. See *Nature* **360**, 219 (1992).

In collaboration with Henri Chrétien of the Nice Observatory in France, Ritchey developed the basic principles of what is now called the Ritchey-Chrétien telescope, a coma-free system of mirrors that has influenced the design of essentially all large research telescopes since the middle of this century.

Hale regarded himself, quite correctly, as a leading international statesman of science. In his lifetime, he was awarded many honours, including membership in leading academies of science around the world. He continues to be widely recognized for his achievements.

Hale did not, however, permit Ritchey to share in the glory of their achievements. Again and again, when Ritchey was about to receive recognition from the astronomical community, Hale intervened to obstruct his nomination. The two astronomers became bitter enemies, each attributing the same faults to the other. The physical and psychological well-being of both men suffered from the resulting stress. Walter S. Adams, Hale's lieutenant and successor as director of Mount Wilson, strongly disliked Ritchey and continually worked against him. At Adams's insistence, Ritchey was denied the use of the 100-inch reflector he had begun and was dismissed without a pension. On Ritchey's death in 1945, only the local weekly newspaper took notice; no daily newspaper carried an obituary.

Donald Osterbrock is eminently qualified to tell the story of Ritchey, Hale and the giant American telescopes. Osterbrock served as director of Lick Observatory from 1973 to 1981. As a research astronomer, he has been intimately acquainted with these instruments and many of the scientists using them. Moreover, he has written extensively on the history of astronomy.

Osterbrock has researched the material for this book with great care. The footnoting is thorough but unobtrusive, referencing hundreds of personal letters, interviews and other sources. Osterbrock puts landmark astronomical research achievements into perspective and traces the evolving concepts within and between generations of astronomers. The book contains pictures of many famous astronomers of the period and an informative appendix on various reflecting telescope configurations.

For any reader interested in the history of modern astronomy or in the development of modern telescopes, I can highly recommend this book. But it is much more than an account of the developing technology; it is a poignant story of human genius and human frailty, and their influence on crucial scientific decisions. □

Joel A. Gwinn is in the Department of Physics, University of Louisville, Louisville, Kentucky 40292, USA.

Neuroscience Titles from Portland Press

Neuropeptide Gene Expression

Edited by A J Turner
Spring 1994 300 pages
I 85578 044 5 H £50.00/\$75.00+

Neurochemistry of Drug Dependence: From Molecular Targets to Behavioural Consequences

Edited by S Wonnacott and G G Lunt
December 1993 200 pages
I 85578 034 8 H £45.00/\$81.00+

Nitric Oxide: Brain and Immune System

Edited by S Moncada, G Nisticò and E A Higgs
December 1993 304 pages
I 85578 046 1 H £78.00/\$124.00+

Proteolysis and Protein Turnover

Edited by J S Bond and A J Barrett
December 1993 250 pages
I 85578 039 9 H £65.00/\$100.00+

Postgraduate Study in the Biological Sciences: A Researcher's Companion

R J Beynon
July 1993 150 pages
I 85578 009 7 P £9.50/\$17.00

The Annexins

Edited by S E Moss
"... this volume is certainly the definitive source of information on this family of proteins."

Trends in Cell Biology
1992 184 pages
I 85578 008 9 H £39.50/\$70.00

Orders to: **Portland Press**
Commerce Way, Colchester
CO2 8HP, UK
Tel 0206 796351 Fax 0206 799331
In USA & Canada: **Portland Press Inc**, Ashgate Publishing Co
Old Post Road, Brookfield,
VT 05036 - 9704, USA
Tel 802 276 3162 Fax 802 276 3837
† Provisional prices



**PORTLAND
PRESS**

Reader Service No.14

The rise and fall of the 17-keV neutrino

Douglas R. O. Morrison

Experiments showing evidence for a heavy neutrino with a mass of 17 keV launched the new particle on an erratic eight-year career, during which it raised questions about the Standard Model of particle physics and about cosmological theories, stimulated many theoretical papers and pushed experimental techniques to their limit. Its demise provides grounds for faith in the efficacy of the scientific method.

IN 1985 John Simpson announced¹ the discovery of a neutrino with a mass of 17 keV. As massive neutrinos had long been anticipated, the announcement caused considerable excitement within the particle-physics community. But because the existence of so massive a neutrino is in contradiction to the Standard Model of particle physics and poses serious problems in cosmology and astrophysics, the result was highly controversial.

Simpson's result came from the measurement of the energy spectrum of β -particles emitted during decay of tritium. In the early part of this century it was believed that β decay was a two-body process, in which a neutron decays into a proton and an electron (the β -particle). If this were the case, all β -particles should have the same energy; but experimentally they were found to have a broad range of energies. Bohr was led to propose tentatively, that this was an instance in which energy conservation was violated, but in 1932 Pauli suggested instead that a third particle was involved in β decay, accounting for the extra energy. This particle, the neutrino, was postulated to have zero, or near-zero, mass and to interact scarcely at all with other particles.

Experiments at the large electron-positron collider (LEP) at CERN have shown there are precisely three types of normal neutrino—the latest value² for the number of neutrino species is 2.980 ± 0.027 , which is in agreement with theories of nucleosynthesis after the Big Bang. The Standard Model can accommodate three neutrinos which could possibly have a small (non-zero) mass, and great efforts are being made to search for any mass. But a neutrino of 17-keV mass would have to have exotic properties if it is to remain consistent with the LEP result—for example, its coupling (that is, interaction strength) to the Z^0 particle, which mediates the weak force, would have to be much smaller than that of the three known neutrinos.

Astrophysicists would welcome a neutrino with low mass as it could be a major component of the dark matter that is generally believed to make up most of the matter of the Universe. But if there were to exist a stable neutrino with a mass of more than only a few tens of electron volts, it would decouple from the primordial plasma and fill the Universe today. For a neutrino with a mass as large as 17 keV, the neutrino mass density would raise the total mass of the Universe far above the maximum value consistent with astronomical observations and cosmological theories.

This article traces the history of the 17-keV neutrino from Simpson's initial report to the current consensus that it does not exist. Following Simpson's announcement¹ a first generation of experiments designed to confirm his findings, gave null results. Simpson criticized these experiments, but interest was revived only when further experiments from several laboratories did appear to support the existence of the 17-keV neutrino. These results spawned many theoretical papers. A second generation of experiments was then designed to search for the particle with greater accuracy and reliability. These experiments were often conducted at the limit of the available technology; all gave null results. The ups and downs of this history are reflected in Fig. 1, which shows the time distribution of papers concerned with the 17-keV neutrino. The excitement of the 1991–92 period has now

abated, and most researchers believe that there is no 17-keV neutrino.

The rise: 1985–91

Simpson¹ studied the electron spectrum from tritium decay, which produces electrons of maximum energy 18.6 keV. His experiment was simple; a solid-state silicon detector was used to detect electrons emitted at all angles from a thin tritium source.

Simpson observed a kink in the otherwise smooth spectrum at 1.5 keV (Fig. 2), which he interpreted by suggesting that two neutrinos were emitted in the β decay: the normal electron neutrino of zero mass, giving a spectrum with an end-point of 18.6 keV, and a neutrino of 17.1-keV mass, giving a spectrum with an end-point of 1.5 keV. This would reproduce the kink observed at 1.5 keV. He interpreted the data as giving a mixing probability (essentially the fraction produced) of 3% for the 17-keV neutrino.

But several other groups^{3–7} soon reported that they did not observe a kink in their own tritium decay spectra. Furthermore, it was pointed out^{8–10} that atomic effects (involving screening, mean excitation energy, neutralization energy and chemical binding) are important for tritium in the low-energy region below 1.5 keV, and could account for most, if not all, of the effect observed. Simpson¹¹ subsequently reduced his estimate of the mixing probability to 1%. After further null experiments^{12,13}, the excitement quickly died down.

But Simpson criticized the null experiments on two grounds. (1) His experiment had been performed without a magnetic field, whereas all of the null experiments except that of Ohi *et al.*⁴, had been conducted in a magnetic field. Simpson noted that the spectra had to be corrected in each case by a 'shape-correction factor' which was required to bring the observed spectrum into agreement with the shape of the theoretical spectrum. He

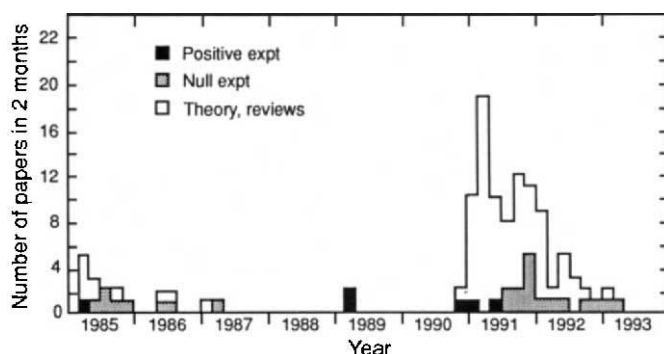


FIG. 1 Arrival times of papers concerned with 17 keV neutrinos. Experimental papers with positive results are shown as black, null papers (finding no evidence for a 17 keV neutrino) are shaded. Three changes from positive to null results are counted as null papers. This compilation is based on preprints received by the CERN library and the author's collection, so it may be missing some papers, especially theoretical ones. The time chosen is as close as possible to the date of first submission or first presentation, which is, generally not the date of publication. Updates and duplicates have been excluded.

asserted that these shape-correction factors could conceal the presence of a kink in the β spectra. (2) Simpson asserted¹¹ that, in the published spectra of some of the null experiments, there were in fact kinks corresponding to a 17-keV neutrino. For the important case of Ohi *et al.*⁴, who used no magnetic field, this was done by selecting an asymmetric range around a fluctuation which was taken as a 'kink' and then rotating the base line even though this gave an unacceptable end-point (see Fig. 6 and 7 of ref. 14). Similarly the claim that Apel'kov *et al.*⁵ had missed a kink is rather optimistic (see Figs 2c and 8 of ref. 14). Despite these comments, it was generally considered that the claim was dead.

But in 1989 the 17-keV neutrino was revived when Simpson and Hime performed new experiments^{15,16} on the decay of both tritium and ³⁵S. The latter is important as the end-point of the ³⁵S spectrum is at 165 keV, so the kink would be predicted at 150 keV—far from the region in which atomic effects are important. They found kinks consistent with a 17-keV neutrino with a mixing ratio of about 1%. The earlier null experiments were again criticized. These results failed, however, to stimulate further theoretical or experimental activity.

But in late 1990 and early 1991, further experiments provided this stimulus. Hime and Jelley¹⁷ performed an improved version of the experiment of Simpson and Hime. Back-scattering of electrons from the target poses a problem in this experiment. Simpson and Hime¹⁵ indicated that, if the incident electrons were completely diffuse, 30% were back-scattered—they stated that their experiment resembled that of Planskoy¹⁸ who measured a back-scattering fraction of 32%. If, in contrast, electrons were incident normal to the surface only, the fraction back-scattered was only 15%. So Hime and Jelley introduced collimators in front of the source and before the detector, and also an intermediate baffle which they claimed would exclude stray scattered electrons (Fig. 3). For a ³⁵S source, they found evidence for a neutrino with a mass of 17.0 ± 0.4 keV and a mixing ratio of $0.84 \pm 0.08\%$. This is generally considered the most significant evidence for a 17-keV neutrino. Later Hime and Jelley¹⁹ reported that ⁶³Ni gave similar results, but with lower statistics. Sur *et al.*²⁰ used a ¹⁴C source and found evidence for a neutrino of mass $17.1 \pm \text{keV}$ with a mixing ratio of $1.2 \pm 0.3\%$; Zlimen *et al.*²¹ also found weak evidence (about two standard deviations) for a 17-keV neutrino.

These results for the mass of the putative neutrino are collected in Fig. 4. All of these experiments used solid-state detectors or inner bremsstrahlung and no magnetic spectrometers. Six of the seven experiments indicated masses very close to 17 keV; DiGregorio *et al.*²² found a mass of 13.8 ± 1.8 keV with an upper

limit of 0.5% mixing (95% confidence limit, CL). Taken in isolation, these results looked impressive, and they inspired many further experiments and a spectacular increase in the number of theory papers. Glashow²³ attempted to explore the cosmological implications of a 17-keV neutrino as well as the possible relevance to the solar-neutrino problem⁴⁴. To solve the problem of an excessively massive universe, Simpson²⁴ suggested that the 17-keV neutrinos are unstable. He pointed out that there was a gap in the observed time distribution of neutrinos from supernova 1987A, which he interpreted as indicating that there were two types of neutrino—the early signal was due to 'normal' ones, whereas the later signal resulted from the in-flight decay of 17-keV neutrinos. In this way he deduced a mean life of $0.6 - 1.6 \times 10^4$ s. But Salati²⁵ showed that red-giant cooling theory, together with observations from the Solar Maximum Mission and the Cosmic Background Explorer satellites constrain the decay branching ratio into a lighter neutrino and a photon to be $\lesssim 4 \times 10^{-5}$; that is, very few heavy neutrinos would decay into 'normal' lighter ones. Furthermore, although there are gaps in the time distributions of SN1987A neutrinos measured by the Kamiokande and IMB detectors, there is no such gap when the two time distributions are combined (see Fig. 1 of ref. 26, for example). In any case, the fact that Simpson's model²⁴ actually predicts that there should be no gap in the time distribution after all would seem to invalidate his original hypothesis.

Second-generation experiments

In December 1991 a three-day meeting on the 17-keV neutrino was held in Berkeley²⁷. At that stage, 17 experiments were either in progress or had been proposed. The critical issues that these experiments had to address were as follows. (1) How could one cope with the problem of back-scattering of electrons from the detector? The fraction back-scattered varied between 15 and 30%, depending on the incident angle, which is much greater than the effect searched for, and very much greater than the error of 0.1% claimed¹⁷. Furthermore this fraction varied slightly with the energy of the electron. (2) If the upper limit to the mixing were much less than 0.8%, could one be sure that the experiment would detect an effect even if it were present? (3) Can such an experiment be done without using a magnetic spectrometer, which introduces important shape factors? (4) If there were no 17-keV neutrino, how could one explain the positive results and in particular why they agreed on the value of 17 keV?

None of the second-generation experiments reported so far

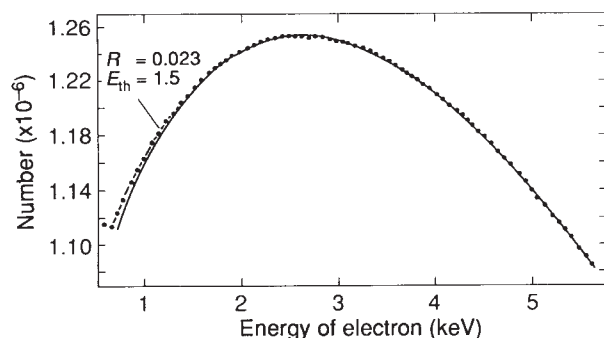


FIG. 2 Part of the distribution of energies of electrons from the β decay of tritium reported by Simpson¹. The smooth curve is the expected spectrum with only a massless neutrino produced, while the dashed curve shows the expected effect of adding a massive neutrino of 17.1 keV, with a coupling R , of 2.3% rather than the 3% given in the text. The next point to the left of the lowest-energy point is off-scale at 1.6×10^6 counts. E_k is the value in keV at the kink.

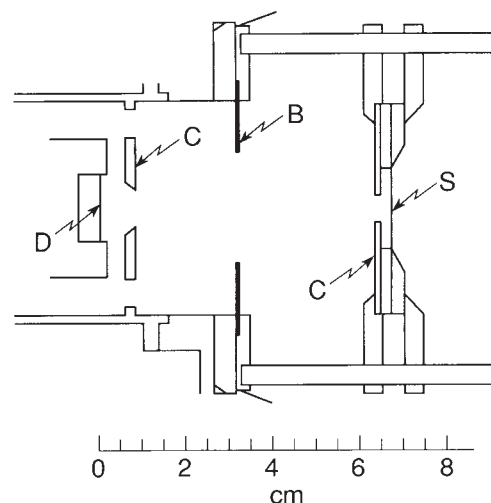


FIG. 3 Inner part of the experimental apparatus used by Hime and Jelley^{17,19}. B is the annular baffle that caused problems, S is the source from which the decay electrons emerge, D is the solid-state detector and C represents the two collimators around the source and the detector.

have found evidence for a 17-keV neutrino (at the level of 0.8% mixing). The two most significant, conducted in mid-1992, were those of Oshima *et al.*^{28,29} (carried out at INS, Tokyo) and Mortara *et al.*³⁰ (at Argonne). Oshima *et al.*^{28,29} used a magnetic spectrometer but had such high statistics (Fig. 5) that the shape factor was unimportant. They found a mixing strength of $-0.0011 \pm 0.033 \pm 0.030\%$ (the first error being statistical and the second systematic) and an upper limit for a 17-keV neutrino of $<0.073\%$ at 95% CL. They commented that it has been known since 1938 that experimental spectra do not quite have the form predicted theoretically, and that it was realized in the 1940s that this was due to energy losses in the source and back-scattering.

The Argonne experiment³⁰ (Fig. 6) was very ingenious. No collimators were needed and no shape correction factors were required. They found a mixing probability of $-0.04 \pm 0.08 \pm 0.08\%$, with no evidence for a 17-keV neutrino at an upper limit of 0.25% mixing (95% CL). To test whether the experiment had sufficient sensitivity they added to the ^{35}S source a weak ^{14}C source (1.34% of the activity), which was expected to produce a kink about 11 keV below the end-point. The results indicated $1.4 \pm 0.1\%$ mixing from the ^{14}C source. (Chen *et al.*³¹ had previously 'calibrated' their spectrum along the same lines, by covering part of their source with a thin absorber.) This experiment answered all the objections raised for the other null experiments, and led Hime³² to confess that he was "baffled". This was interpreted by many as meaning that the existence of the 17-keV neutrino had been disproved.

Several other null results were reported. The Princeton group³³ extended their earlier experiment using a magnetic spectrometer and a ^{35}S source. They had better detectors, improved data analysis and a much larger data set, and their shape factor was flat. They found a mixing ratio of $0.1 \pm 0.15\%$, or zero within errors. The Berkeley group, which had previously reported evidence for the 17-keV neutrino, performed a different experiment³⁴ looking at the internal bremsstrahlung photon spectrum from the electron capture decay of ^{55}Fe . They could exclude the existence of a 17-keV neutrino with a mixing of 0.8% by seven standard deviations. They concluded that "the effect reported previously in a lower statistic experiment, and interpreted to be the result of a 17 keV neutrino, is not in fact, caused by a massive neutrino". Experiments with tritium, Simpson's original source, have been repeated by Kalbfleisch and Bahren³⁵ Holzschuh and Kundig³⁶, and Stoeffl *et al.*³⁷, all of whom found no evidence for a 17-keV neutrino with 0.8% mixing.

Experimental conclusions

Jelley (ref. 38 and personal communication) has reported that new experiments have been conducted and are still in progress at Oxford. A variety of source and detector geometries were used and evidence for additional scattering in the apparatus was found. In one run they changed the baffle but kept the same geometry, detector and source and did not find agreement with the earlier positive results. From the fits to their data, they have concluded that, although they do not yet have a complete understanding of the electron response function, the original data cannot now be regarded as providing positive evidence for a 17-keV neutrino.

Hime^{39,40} has analysed again the effects of the intermediate baffle shown in Fig. 3, as well as effects from the apertures and the source backing. Recalculating the three runs with these effects, similar χ^2 fits could be obtained either by assuming a 17-keV neutrino with $\sim 0.8\%$ mixing or with no extra neutrino but with the new corrections for the baffle, apertures and source backing. For the Oxford experiments with ^{35}S and ^{63}Ni , he now quotes upper limits of 0.36% and 0.53% (90% CL) respectively. The main problem was with the baffle, the edge of which can be seen directly by the source and by the detector and so can scatter electrons directly (this had been pointed out earlier—by Piilonen and Abashian⁴¹, for example). It is unfortunate that the experiment has not been repeated with three or four baffles. Hime

has concluded that he is indeed "baffled" and that "The anomalies observed in β -decay measurements at Oxford can be re-interpreted after account is made for a more complete electron response function". He has also re-examined the earlier experiments^{15,16}, in which no baffle was used, and found that placing a ring around the detector gave similar effects. He suggests that mysteries remain and that further experiments would be useful, but one doubts that many will be made.

Simpson (personal communication) is doing some further experiments to search for effects that could mimic a neutrino kink in β spectra and feels that comments would be premature until these experiments are completed.

Recently the Berkeley group of Norman (personal communication) identified cross-talk between pre-amplifiers in their first ^{14}C experiment. After correcting this, they took new data for three-and-a-half months, and one month of background, and found no evidence for a 17-keV neutrino.

Thus the main positive results in favour of a 17-keV neutrino have disappeared, and many high-quality experiments have found no evidence for it with very low mixing limits; it is generally considered that the experimental evidence argues against its existence.

Conclusions and lessons

Initially there was a consistent over-enthusiasm to promote the 17-keV neutrino and to find fault with null experiments. The suggestion that experiments with magnetic spectrometers are inferior to those without a magnetic field surprised many, as experience has shown that adding a magnetic field can help to avoid artefacts. It was correct and useful for Simpson to point out that shape factors have to be employed, and that they can

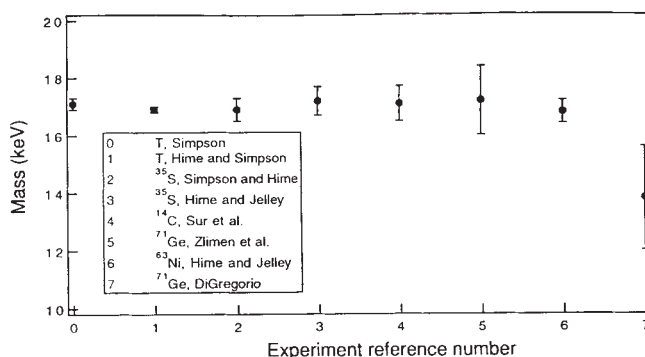


FIG. 4 Values of the mass of a heavy neutrino from positive determinations, from ref. 43.

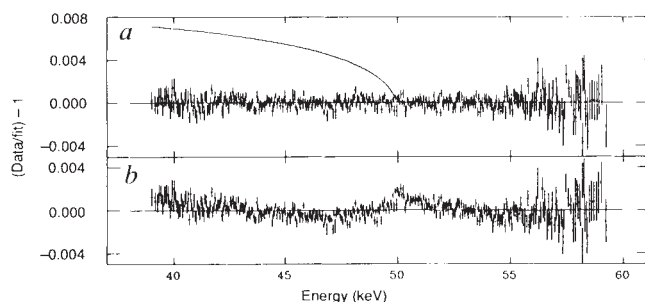


FIG. 5 Data from the experiment of Kawakami *et al.*^{28,29}. They have 2.4×10^9 events in 1,800 bins around 50 keV, the expected kink position for ^{63}Ni . a, Plotted relative to a best global fit with the mixing as a free parameter and b, with a fixed 1% mixing. The curve in a illustrates the size of a 1% mixing effect of a 17 keV neutrino.

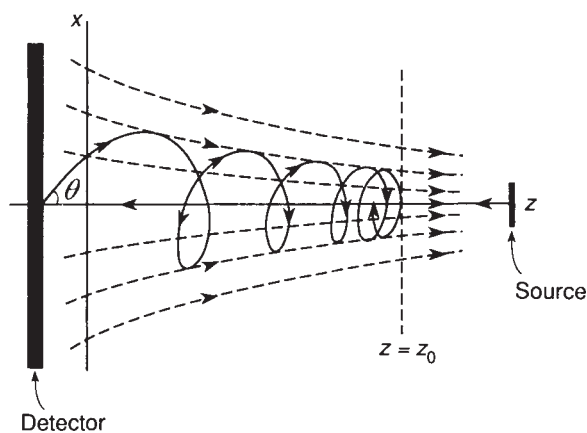


FIG. 6 Illustration of the magnetic mirror used in the experiment of Mortara *et al.*³⁰. It has a solenoid magnetic field which peaked at the ³⁵S source (at the centre) and was lower at the cooled Si(Li) detector (in the fringe field). The lines of magnetic field are shown dashed. An electron is shown leaving the source, hitting the detector and being back-scattered at an angle θ , spiralling away in the magnetic field. For example, an electron leaving the source at 80° to the normal, was then incident on the detector at 30° reducing the effect of back-scattering which is strongly angle dependent. Also any electrons that were back-scattered from the detector would spiral back towards the source due to the high field. z and x are the horizontal and vertical axes; the maximum distance an electron can travel when back-scattered from the detector, is to the plane $z = z_0$.

be of the order of several per cent, which is more than the mixing claimed—this argument was supported later by Monte Carlo calculations of Bonvicini⁴². It would be surprising, however, if many null experiments, using different equipment, all had shape factors that concealed a kink. De Rujula has jokingly noted that, as the 17-keV neutrino was only observed in the absence of magnetic fields, it must be magnetophobic.

This over-enthusiasm and the selection of only positive experiments, is reminiscent of the Anglo-Saxon legal method, which

it is interesting to compare with the scientific method. In the former, two lawyers oppose one another; each presents the arguments favourable to his side and, where possible, criticizes the other sides' arguments. Good arguments by the opposition are ignored. In the scientific method, each individual is expected to consider all the arguments, theories and facts and to evaluate them critically to obtain a consistent conclusion, even though it may require one to reject some theories and/or experiments.

The major lesson of the 17-keV neutrino story is that the scientific method triumphed. It is normal and to be expected that many experiments will produce a few wrong results—the problem is to detect them. Neutrinos are exceptional particles, having no charge, no (or almost no) mass and very low reaction rates, so it is not surprising that the history of neutrino physics has been afflicted with many results that needed correcting later. In addition the effects observed in this case were small, the errors very small and the analysis sensitive to these errors⁴².

When results supporting the 17-keV neutrino became stronger, further experiments were designed carefully to avoid the ambiguities. Considerable ingenuity was employed and new techniques were developed which may generate useful spin-offs. Greater understanding has been gained about the use of solid-state detectors and magnetic spectrometers.

It is usual that wrong results produce a few people who continue to believe in them. Blondlot, for example, published a book on N-rays two years after they were shown not to exist by Wood. And although the consensus of the scientific community has now turned against cold fusion, a few hundred researchers continue to list positive results, whether consistent or not, and to ignore or criticize the more numerous null experiments often performed with superior apparatus. In contrast, there are some who listen to criticism and who are self-critical, who take the possibility of error seriously and perform experiments designed to look for falsification. It is to the credit of Andrew Hime, Nick Jelley and Eric Norman that they have resisted the human desire not to prove oneself wrong, have checked and repeated their experiments and have announced errors that are identified. One respects them for following the tradition of the scientific method. □

D. R. O. Morrison is at CERN, CH-1211 Geneva, Switzerland.

1. Simpson, J. J. *Phys. Rev. Lett.* **54**, 1891–1893 (1985).
2. Combined results of the four LEP experiments—ALEPH, DELPHI, L3 and OPAL—presented by J. Lefrançois at *Int. Europhysics Conf. on High Energy Phys.* Marseille, 22–28 July 1993.
3. Alizotoglou, T. *et al. Phys. Rev. Lett.* **55**, 799–802 (1985).
4. Ohi, T. *et al. Phys. Lett.* **B160**, 322–324 (1985).
5. Apatikov, A. M. *et al. Zh. eksp. teor. Fiz.* **42**, 233 (1985); also as *JETP Lett.* **42**, 289–292 (1985).
6. Datar, V. M. *et al. Nature* **318**, 547–548 (1985).
7. Markey, J. & Boehm, F. *Phys. Rev.* **C32**, 2215–2216 (1985).
8. Lindhard, J. & Hansen, P. G. *Phys. Rev. Lett.* **57**, 965–967 (1986).
9. Drukarev, E. G. & Strikman, M. I. *Phys. Lett.* **B186**, 1–5 (1987).
10. Koonin, S. E. *Nature* **354**, 468–470 (1991).
11. Simpson, J. J. *Phys. Lett.* **B174**, 113–114 (1986); also in *Proc. Moriond Workshop on Massive Neutrinos in Particle Phys. and Astrophys.* (eds Tran Thanh Van, J. & Fackler, D.) 565–577 (Edns Frontieres, Gif-sur-Yvette, 1986).
12. Hetherington, D. W. *Phys. Rev.* **C36**, 1504–1513 (1987).
13. Borge, M. J. G. *et al. Physica Scripta* **34**, 591–596 (1986).
14. Morrison, D. R. O. in *Joint Lepton-Photon Symposium and Europhysics Conf. on High Energy Physics*, Geneva (eds Hegarty, S., Potter, K. & Quercigh, E.) 599–606 (World Scientific, Singapore, 1991).
15. Simpson, J. J. & Hime, A. *Phys. Rev.* **D39**, 1825–1836 (1989).
16. Hime, A. & Simpson, J. J. *Phys. Rev.* **D39**, 1837–1850 (1989).
17. Hime, A. & Jelley, N. A. *Phys. Lett.* **B257**, 441–449 (1991).
18. Planskoy, B. *Nucl. Instrum. Meth.* **61**, 285–295 (1968).
19. Hime, A. & Jelley, N. A. Oxford University Dept. of Nuclear Physics preprint OUNP-91-21 (1991).
20. Sur, B. *et al. Phys. Rev. Lett.* **66**, 2444–2446 (1991).
21. Zilmen, I., Ljubicic, A., Kaucic, S. & Logan, B. A. *Phys. Rev. Lett.* **67**, 560–563 (1991).
22. DiGregorio, D. E. *et al. Tandar*, Argentina preprint LNY584-L1-SB (1992); also as *Phys. Rev. Lett.* (submitted).
23. Glashow, S. L. *Phys. Lett.* **B256**, 255–257 (1991).
24. Simpson, J. J. *Phys. Lett.* **B269**, 454–457 (1991).
25. Salati, P. Report No. A-418/93 (ENSLAPP, Annecy-le-Vieux, 1993).
26. Morrison, D. R. O. in *Moriond Conf. on Progr. in Atom. Phys., Neutrinos and Gravitation* (eds Chardin, G., Fackler, D. & Tran Thanh Van, J.) 207–216 (Edns Frontieres, Gif-sur-Yvette, 1992); also as *Uzbekhi* **154(4)**, 719–752 (1988).
27. *Workshop on the 17 keV Neutrino* (Convenor Sadoulet, B.) Center for Astrophysics, Univ. California at Berkeley, 18–20 Dec. 1991.
28. Oshima, T. *et al. KEK (Natnl. Lab. for High Energies, Japan) preprint* 92-184 (1992); also as *Phys. Rev. D* (submitted).
29. Kawakami, H. *et al. Phys. Lett.* **B287**, 45–50 (1992).
30. Mortara, J. L. *et al. Phys. Rev. Lett.* **70**, 394–397 (1993).
31. Chen, M. *et al. Phys. Rev. Lett.* **69**, 3151–3154 (1992).
32. Hime, A. presented at *XV Int. Conf. on Neutrino Phys. and Astrophys.* Grenada, 1992; *Nucl. Phys. (Proc. Suppl.)* **31**, 50–56 (1993).
33. Berman, G. E., Pitt, M. L., Calaprice, F. P. & Lowry, M. M. *Phys. Rev. Lett.* (in the press).
34. Wietfeldt, F. E. *et al. Phys. Rev. Lett.* **70**, 1759–1762 (1993).
35. Kalbfleisch, G. R. & Bahrman, M. Y. *Phys. Lett.* **B291**, 336–340 (1992).
36. Holzschuh, E. & Kundig, W. presented at *Moriond Conf. on Perspectives in Neutrinos, Atomic Phys. and Gravitation Villars-sur-Ollon*, (1993).
37. Stoeffl, *et al.* reported by Caldwell, D. O. in *Am. Phys. Soc. Div. of Particles and Fields Meeting*, Fermilab, 1992.
38. Jelley, N. presented at *Moriond Conf. on Perspectives in Neutrinos, Atomic Phys. and Gravitation Villars-sur-Ollon*, 31 Jan.–6 Feb. 1993.
39. Hime, A. *Phys. Lett.* **B299**, 165–173 (1993).
40. Hime, A. *Moriond Conf. on Perspectives in Neutrinos, Atomic Phys. and Gravitation Villars-sur-Ollon* (1993).
41. Piilonen, L. & Abashian, A. in *Moriond Conf. on Progress in Atomic Phys. Neutrinos and Gravitation* (eds Chardin, G., Fackler, D. & Tran Thanh Van, J.) 225–234 (Edns Frontieres, Gif-sur-Yvette, 1992).
42. Bonvicini, G. Z. *Phys. A345*, 97 (1993).
43. Robertson, R. G. H. presented at *Int. Conf. on High Energy Phys.* Dallas, August 1993.
44. Bahcall, J. N., Davis, R. Jr & Wolfenstein, L. *Nature* **334**, 487–494 (1988).

ACKNOWLEDGEMENTS. Thanks are due to many people who contributed or commented on this manuscript, particularly F. Boehm, G. Bonvicini, D. O. Caldwell, D. Cline, P. Darriulat, A. De Rujula, R. L. Graham, P. G. Hansen, D. Hetherington, A. Hime, M. Jacob, N. Jelley, E. Kolb, V. A. Lyubimov, E. Norman, T. Oshima, L. Piilonen and T. Yamazaki. The opinions expressed in this article are those of the author and not necessarily shared by the above. D. Dallman is thanked for compiling preprints received by the CERN library.

The DNA replication fork can pass RNA polymerase without displacing the nascent transcript

Bin Liu^{*}, Mei Lie Wong^{*}, Rachel L. Tinker[†], E. Peter Geiduschek[†]
& Bruce M. Alberts^{*‡}

^{*} Department of Biochemistry & Biophysics, University of California, San Francisco, California 94143-0448, USA

[†] Department of Biology, University of California, San Diego, La Jolla, California 92093-0634, USA

Replication proteins encoded by bacteriophage T4 generate DNA replication forks that can pass a molecule of *Escherichia coli* RNA polymerase moving in the same direction as the fork *in vitro*. The RNA polymerase ternary transcription complex remains bound to the DNA and retains a transcription bubble after the fork passes. The by-passed ternary complex can resume faithful RNA synthesis, suggesting that the multisubunit RNA polymerase of *E. coli* has evolved to retain its transcript after DNA replication, allowing partially completed transcripts to be elongated into full-length RNA molecules.

No known mechanism prevents DNA replication and transcription from taking place on a DNA molecule concurrently; when they move in the same direction, the respective polymerases must use the same DNA single strand as template. In *E. coli*, the rate of replication is 10–15 times faster than the rate of transcription^{1,2}, so that collisions between the two types of polymerase are inevitable, even when they move in the same direction. Collisions could be resolved in one of three ways: a replication fork could knock an RNA polymerase molecule and its nascent transcript out of its way; a replication fork could slow down and passively follow behind a transcription complex; or a replication fork could pass a transcribing RNA polymerase molecule without displacing it from the template.

Preservation of the nascent transcript when a replication fork passes would be advantageous because RNA chains are generated by an energy-consuming multistep process^{3–5}. Is this possible chemically? To answer this question, we examined the consequences of a collision between a replication fork and co-directionally transcribing RNA polymerase. We used the highly purified *in vitro* T4 bacteriophage DNA replication system to replicate through a precisely placed *E. coli* RNA polymerase transcription complex. In this completely defined system, the fate of the nascent transcript after replication can be determined unambiguously. Surprisingly, a replication fork can pass through a transcription complex without displacing it, leaving intact its ability to resume RNA chain elongation.

A template for investigating the collision

A uniquely nicked circular DNA molecule containing an appropriately oriented strong T4 late promoter^{6–8} was used as a DNA template that supports co-directional replication and transcription (Fig. 1a, left side; note that the template strand for transcription is also the template for leading-strand DNA synthesis). By withholding rCTP, we stalled the RNA polymerase at a specific downstream site, creating a stable ternary transcription complex composed of RNA polymerase, an 18-nucleotide (nt) nascent RNA transcript, and the DNA template. The core RNA polymerase was from *E. coli*; the T4 gene 55 σ -family protein enables it to recognize the T4 late promoter⁹.

To reduce the binding of RNA polymerase to the nick¹⁰ and

to weak variant T4 late promoters on the plasmid, we either used a low molar ratio of RNA polymerase to template DNA (for example, 4 : 1), or purified the ternary complex on Sepharose CL-2B after exposure to high salt (0.5 M NaCl), as specified in each experiment. The Sepharose CL-2B column excludes the ternary complex, but includes free core RNA polymerase, gene 55 protein and nucleotides. Promoter-bound RNA polymerase and other, less stable ternary complexes dissociate from DNA in 0.5 M NaCl (ref. 11). Thus, the desired ternary complex is highly enriched after passage through the CL-2B column. Moreover, because the gene 55 protein and ribonucleoside triphosphates (rNTPs) have been removed, there is no reinitiation by RNA polymerase during the subsequent DNA replication reaction.

The replication fork passes the ternary complex

Seven highly purified bacteriophage T4-encoded proteins reconstitute an *in vitro* replication system that catalyses efficient leading strand DNA synthesis^{12–15}. The proteins involved are the T4 DNA polymerase holoenzyme (consisting of the products of T4 genes 43, 44, 62 and 45), a helix-destabilizing single-stranded DNA-binding protein (gene 32 protein), the highly processive DNA helicase (gene 41 protein), and the gene 59 protein that greatly facilitates the loading of the gene 41 protein onto DNA at a replication fork (J. Barry and B.M.A., manuscript in preparation).

We analysed the effect of stalled RNA polymerase ternary complexes on the movement of replication forks by alkaline agarose gel electrophoresis. As the DNA template, we used either mock-treated DNA, or CL-2B-purified ternary complexes. Even though about 70–80% of the DNA molecules bear a bound ternary complex (determined by a gel shift assay), there is no strong blockage of DNA synthesis, with or without DNA helicase (gene 41 protein) (Fig. 1b). Thus, the ability to pass the RNA polymerase ternary complex is intrinsic to the DNA polymerase holoenzyme (DNA polymerase plus accessory proteins). When helicase is included in the reaction, the replication fork speeds up, and it advances at a slightly reduced rate on templates bearing the ternary complex (compare lanes 7, 8 with lanes 11, 12), suggesting that the fork pauses transiently before passing stalled RNA polymerase. Without a helicase, the fork pauses at many sites, making it difficult to detect any additional pausing

[‡] To whom correspondence should be addressed.

caused by the RNA polymerase ternary complex (compare lanes 1–3 with lanes 4–6).

The complex stays bound to DNA

The experiment in Fig. 1b shows the DNA replication fork readily passing a DNA template-bound RNA polymerase molecule that carries a transcript. To distinguish between the possible fates of this RNA polymerase (Fig. 1a), we designed the experiment illustrated in Fig. 2a. We used RNA-labelled ternary complexes as templates for replication with dUTP as one of the four dNTP substrates. DNA containing dUMP on one strand is resistant to double-strand cleavage by the restriction enzyme *DraI*, which recognizes the sequence TTAAA. The sensitivity of the RNA-labelled replication products to *DraI*, as analysed by non-denaturing polyacrylamide gel electrophoresis, can therefore be used to analyse whether the replication fork has passed the ternary complex without displacing it (see Fig. 2a).

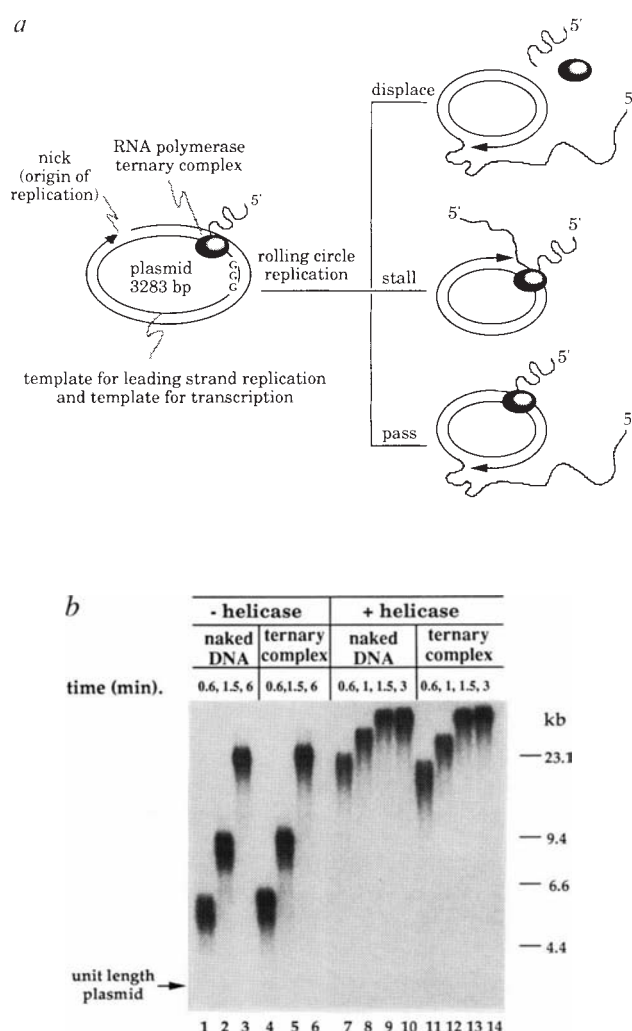
The analysis of such an experiment is shown in Fig. 2b (replication with DNA helicase) and Fig. 2c (replication without DNA helicase). Because the same amount of RNA-labelled ternary complex is seen in lane 1 (no replication) and lane 2 (after

FIG. 1 a, The experimental system. The template for *in vitro* replication by the bacteriophage T4 replication proteins is a 3.3-kilobase-pair (kb) circular plasmid containing the replication origin of bacteriophage M13, located ~170 nt behind the stalled RNA polymerase. Nicking this origin with the filamentous bacteriophage gene 2 endonuclease provides a unique DNA 3' end that serves as a starting site for initiation of rolling circle DNA synthesis *in vitro*³⁰. Three consecutive G nucleotides were placed on the template strand 17, 18 and 19 base pairs (bp) downstream of the transcription initiation site. Using the dinucleotide UpG to initiate transcription at bp -1 (that is, one bp upstream of the normally initiating G) in the presence of rATP, rGTP and rUTP and withholding rCTP, we stalled the RNA polymerase at the triple G site with an 18-nt nascent transcript. Because rCTP is withheld, lagging strand DNA synthesis is very inefficient, and we have generally omitted the DNA primase (gene 61 protein)^{36,37} from replication reactions, leaving the template for lagging strand DNA synthesis as a displaced single strand. b, Effect of the ternary complex on movement of the replication fork. The products of *in vitro* DNA synthesis, using either naked DNA (as control) or column-purified ternary complexes as the DNA template, were analysed by alkaline agarose gel electrophoresis, followed by autoradiography. METHODS. a, The plasmid pRT510-C+18, which is derived from pDH310 (ref. 38) through two rounds of mutagenesis³⁹, contains a -35 σ^{70} consensus sequence placed upstream of the -10 T4 late promoter consensus sequence of gene 23. The resulting promoter, P'23, is efficiently used *in vitro* by both σ^{70} -containing and gene 55 protein-containing RNA polymerases. The P'23 sequence was further changed to -1GATATGAAGAGTTGGATCCC, where +1 designates the start site of transcription (non-template strand; the entire sequence of plasmid pRT510-C+18 is available on request). To initiate DNA synthesis on circular pRT510-C+18, the DNA was specifically nicked at the M13 bacteriophage gene 2 protein recognition site, as described⁴⁰. To prepare the ternary complex, 0.2 pmol nicked DNA was incubated with the following reagents in 40 μ l for 30 min at 37 °C: 6 pmol *E. coli* RNA polymerase core, 30 pmol gene 55 protein, 9 pmol gene 33 protein, 27 pmol gene 44/62 protein, 98 pmol gene 45 protein, 1 mM dATP; 100 μ M UpG, 4 μ M rATP, 4 μ M rGTP, 4 μ M [α -³²P]rUTP (specific activity ~50,000–100,000 c.p.m. pmol⁻¹), 33 mM Tris-acetate (pH 7.8), 250 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM dithiothreitol (DTT), 100 μ g ml⁻¹ nuclease-free BSA as protein carrier. The reaction was stopped by chilling the sample on ice, followed by the addition of NaCl to 0.5 M and gel filtration through a 1 ml CL-2B column with a 200- μ l 0.5 M NaCl loading zone, and elution with replication buffer (33 mM Tris-acetate (pH 7.8), 66 mM potassium acetate, 10 mM magnesium acetate, 100 μ g ml⁻¹ BSA and 0.5 mM DTT) in the presence of 3–5% glycerol. Radioactive fractions were pooled for the subsequent replication reactions. Typically 70–80% of the DNA templates were occupied by a ternary complex (determined by gel shift assay). b, Replication in 40 μ l of the replication buffer with 0.02 pmol of the column-purified ternary complex or control naked DNA, 3 μ g ml⁻¹ gene 43 protein, 80 μ g ml⁻¹ gene 32 protein, 40 μ g ml⁻¹ gene 44/62 protein, 20 μ g ml⁻¹ gene 45 protein, 20 μ g ml⁻¹ gene 41 protein and 1.2 μ g ml⁻¹ gene 59 protein (whenever the gene 41 protein was omitted, so

replication), it is evident that the ternary complex is not displaced from the template by the replication fork (quantification of the radioactivity typically shows <5% difference). The slowly migrating, branched structures that would be expected for replication forks stalled behind the ternary complex are not seen. Lane 4 shows the *DraI*-resistant products, proving that the replication fork has passed through the ternary transcription complex. About 30–40% of the DNA in these RNA-labelled complexes is cut by *DraI*, in agreement with other results indicating that 60–70% of the DNA template molecules replicate in this experiment. Lanes 3 in Fig. 2b and c show that *DraI* digestion goes to completion when DNA is replicated with dTTP. Note that the same results are obtained with or without DNA helicase present.

Retention of a transcription bubble

The experiment outlined in Fig. 3a examines whether the ternary complex retains its original position after passage of a replication fork by separately marking DNA for the presence of a ternary complex and for downstream penetration of the replication fork. Enhanced reactivity of the DNA with KMnO₄ (a footprint)



was the gene 59 protein), 25–50 μ g ml⁻¹ rifampicin, 0.5 mM dATP, 0.5 mM dGTP, 0.2 mM dCTP, and 0.08 mM [α -³²P]dTTP (~25,000 c.p.m. pmol⁻¹). After 30 s at 37 °C, non-radioactive dTTP was added to 1 mM to stop the labelling. Aliquots were taken at the indicated times, mixed with Na₃-EDTA (20 mM final concentration), loaded on a 0.6% agarose alkaline denaturing gel, and run in 30 mM NaOH, 1 mM Na₃-EDTA for 18 h at 2 V cm⁻¹.

marks the transcription bubble of the ternary complex^{5,16}; incorporating 5'-methyl dCMP (dmCMP) in place of dCMP into the newly synthesized DNA generates resistance to cutting by the restriction enzyme *AluI*. Any *AluI*-resistant DNA that retains the KMnO₄ footprint of the ternary complex can only be generated by replication forks that have replicated past the ternary complex without permanently displacing this complex.

An analysis of the ternary complex footprint by primer extension is shown in Fig. 3b. A comparison of lanes 2 and 5 reveals no significant reduction of the footprint signal after replication (typically <5% difference). Proof that the ternary complex is not displaced from the template after replication comes from the demonstration (lane 4) that 40–50% of the molecules that carry a ternary complex also resist *AluI* cutting (and therefore must have replicated). When dCMP instead of dmCMP is incorporated into DNA, *AluI* is fully active and the footprint disappears as expected (lane 3). Because the position of the footprint is unchanged after replication (lanes 2, 4 and 5), the by-passed ternary complex retains its place on DNA and its transcription bubble.

A by-passed complex remains fully functional

We next assessed the functional competence of ternary complexes after replication forks have passed through them. Ternary complexes bearing nascent transcripts labelled with [α -³²P]rUTP were purified through CL-2B. Replication proteins were added and replication was allowed to proceed until the fork had travelled several times around the circular DNA template. Non-radioactive rNTPs were then added to permit the elongation of any nascent transcripts. If the ternary complexes are inactivated by the passage of the replication fork, the pre-labelled, 18-nt nascent transcripts should not be elongated into full-length

RNA. No new ternary complexes should form under our experimental conditions (no rNTPs or gene 55 protein present during replication; no gene 55 protein present during the chase); moreover, any newly initiated transcripts would not be radioactively labelled.

The results of the above experiments are shown in Fig. 4a. Lane 1 shows the expected 18-nt nascent transcript before a chase. Lane 2 shows that, as expected, the nascent transcripts on column-purified ternary complexes chase into 427-nt full-length RNA in the absence of DNA replication. The important result is that the 18-nt transcripts are also nearly completely converted to full-length transcripts following replication without or with DNA helicase (lanes 4 and 6, respectively). When no rNTPs are added after replication, a 'mock' 6–8 min incubation leaves the nascent 18-nt transcript unchanged (lanes 3 and 5).

To assess the fidelity of RNA synthesis after replication, we repeated the chase experiment on a DNA template cut with *AluI* to generate only a ~33-nt run-off transcript. Identical run-off transcripts were observed before and after replication (Fig. 4b), demonstrating the precise retention of position by the functional ternary complex.

To test whether a nascent transcript that has been released into solution can reassociate with DNA to be further elongated, we added purified 18-nt ³²P-labelled RNA to a reaction mixture containing the components of the chase experiment shown in Fig. 4a. When incubated with RNA polymerase core (with or without gene 55 protein) and cold rNTPs (either alone or with DNA replication proteins and dNTPs), no 18-nt RNA was elongated, and most of this RNA remained detectable as a radioactive band in the 18-nt position (data not shown).

The above chase experiments are significant if a major fraction of the DNA molecules bearing ternary complexes have been

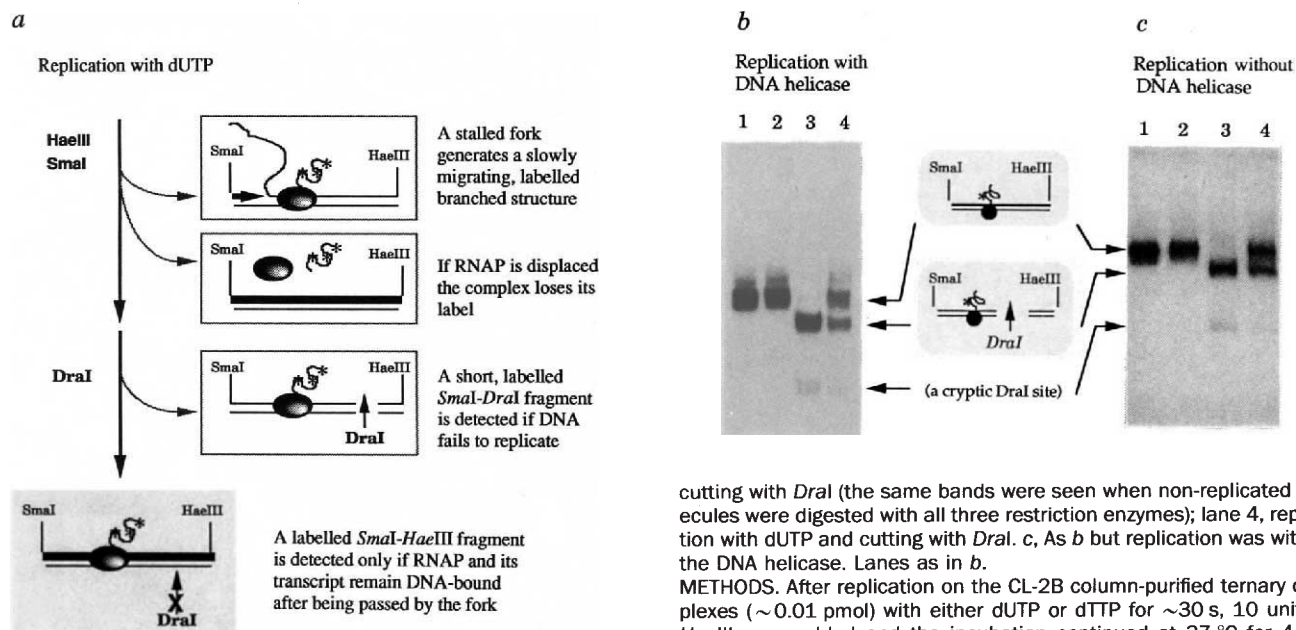


FIG. 2 A test for retention of the RNA polymerase (RNAP) ternary complex, identified by its radioactive nascent transcript, after replication. **a**, Outline of the experiment. After DNA is replicated with dUTP in place of dTTP, a SmaI–HaeIII fragment bearing the ternary complex is tested for its susceptibility to DraI. **b**, Gel autoradiograph after replication through RNAP ternary complexes with DNA helicase (gene 41 protein) present. Lane 1, control ternary complex on the SmaI–HaeIII fragment (no replication); lane 2, control ternary complex on the SmaI–HaeIII fragment after replication with dUTP; lane 3, replication with dTTP and

cutting with DraI (the same bands were seen when non-replicated molecules were digested with all three restriction enzymes); lane 4, replication with dUTP and cutting with DraI. **c**, As **b** but replication was without the DNA helicase. Lanes as in **b**.

METHODS. After replication on the CL-2B column-purified ternary complexes (~0.01 pmol) with either dUTP or dTTP for ~30 s, 10 units of HaeIII were added and the incubation continued at 37 °C for 4 min (without helicase) or 1–2 min (with helicase). The DNA was cut with 10 units of SmaI at room temperature for 5 min. Where indicated, 10 units of DraI were then added for another 5 min at 37 °C. The reaction was stopped by chilling on ice; heparin and Ficoll were added to 100 µg ml⁻¹ and 3%, respectively. Samples were loaded on a 3% (Fig. 2b) or 4% (Fig. 2c) non-denaturing, neutral polyacrylamide gel (37.5:1 acrylamide:bisacrylamide in 1 × TBE (89 mM Tris base, 89 mM boric acid, 2.5 mM EDTA)) for electrophoresis at room temperature for ~5 h at 11 V cm⁻¹. The gel was dried and autoradiographed. *In vitro* replication was done as described in Fig. 1b except that 0.2 mM non-radioactive dTTP or dUTP was used instead of [α -³²P]dTTP.

replicated. To determine this fraction, nascent RNA was labelled with [α - 32 P]rUTP and ternary complexes were purified by gel filtration. Non-replicated circular DNA templates run as a defined band during electrophoresis on neutral agarose gel. Replication converts these molecules to circular molecules with long single-stranded tails, which migrate more slowly. Because only the RNA is labelled, the changing distribution of radioactive signals in the gel as a function of time reflects the efficiency of replication on templates bearing ternary complexes (Fig. 4c). Quantification of radioactivity at the position corresponding to the non-replicated template reveals that $\sim 70\%$ of the templates bearing a ternary complex have been replicated. Moreover, there appears to be no blockage of replication fork movement by the ternary complex, because no discrete bands corresponding to stalled structures appear on the gel, even in the absence of DNA helicase. We conclude that most of our DNA templates have undergone extensive DNA synthesis, and that the ternary transcription complexes bound to them remain functional for RNA chain elongation after the passage of replication forks.

Electron microscopic examination

As an independent test of our conclusions, we have used electron microscopy to examine the fate of the ternary complex after replication. The analysis should also reveal unanticipated products of replication, if any are formed. For each DNA molecule that undergoes rolling-circle replication, the extent of such replication is easily assessed by the length of its single-stranded

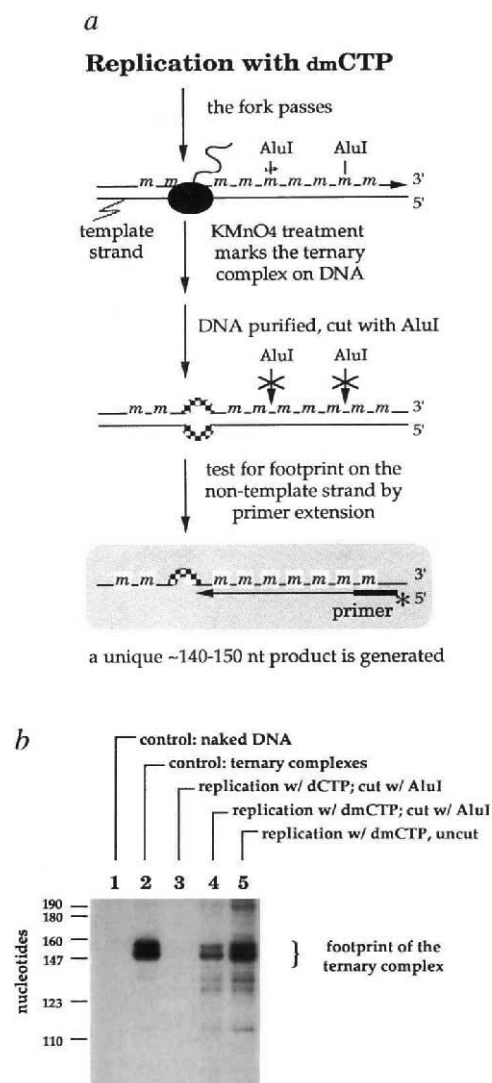
DNA tail. In principle, the replication fork must have passed the ternary complex without displacing it from the template if a DNA molecule bearing such a complex has a single-stranded tail that is longer than the distance from the nick to the ternary complex (~ 170 nt).

The template for these replication reactions was the ternary complex purified on the CL-2B column. Before visualization, replication products were re-treated with 0.5 M NaCl and passed again through CL-2B to remove replication proteins. As a control, Fig. 5a shows a globular particle associated with the non-replicated circular template. Several lines of evidence suggest that this particle is the ternary complex: (1) it survives high-salt (0.5 M NaCl) treatment and CL-2B gel filtration; (2) it occupies the expected place on DNA cut at unique restriction enzyme sites (such as *SspI* in Fig. 5b and *HindIII* in Fig. 5c); (3) it disappears when rNTPs are added for several minutes (not shown); (4) nascent transcripts can be seen on brief incubation (30 s) with a low concentration of rNTPs (1 μ M each) (Fig. 5d).

Replicated DNA molecules (Fig. 5e-g) bearing the ternary complex have tails of varying lengths that can exceed the size of the circular template. We randomly sampled 180 molecules to obtain the data in Fig. 5i, j. A similar fraction of the DNA molecules bear the ternary complex before and after replication (Fig. 5i), consistent with the finding that the by-passed ternary complex remains DNA-bound. Moreover, a significant fraction of templates bearing the ternary complex have tails longer than the size of the circular template (Fig. 5j), proving that at least

FIG. 3 Determination of the location of the ternary complex after replication. a, Outline of the experiment. KMnO_4 oxidizes regions of single-stranded DNA⁴¹ in the ternary complex, and this footprint of the transcription bubble is observable by primer extension analysis only if DNA is resistant to cutting by *AluI* at the sites shown. Resistance is conferred by incorporation of dmCMP ('m'). The asterisk represents the 32 P-label at the 5' end of the primer. b, Primer extension analysis. Lane 1, naked DNA control; lane 2, ternary complex control, showing the position and signal intensity of the ternary complex in unreplicated DNA not cut with *AluI*; lane 3, replication with dCTP and cut with *AluI*; lane 4 replication with dmCTP and cut with *AluI*; Lane 5 replicated with dmCTP, but not cut with *AluI*.

METHODS. Nicked DNA (0.1 pmol) was incubated with 0.4 pmol RNA polymerase core supplemented with 2 pmol gene 55 protein and 0.8 pmol gene 33 protein; 40 μ g ml⁻¹ gene 44/62 protein, 20 μ g ml⁻¹ gene 45 protein, 1 mM dATP, 120 μ M UpG, 5 μ M rATP, rGTP, and rUTP, 220 μ M 3'-O-methyl rCTP as chain terminator, 5% polyethylene glycol (3.3K), 33 mM Tris-acetate (pH 7.8), 250 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT and 100 μ g ml⁻¹ nuclease-free BSA. After incubation at 37 °C for 15 min, potassium acetate was diluted to 120 mM and replication was allowed to proceed at 37 °C for 4 min by adding 3 μ g ml⁻¹ gene 43 protein, 60 μ g ml⁻¹ gene 32 protein, 20 μ g ml⁻¹ gene 41 protein, 0.2 μ g ml⁻¹ gene 59 protein, 0.5 mM dGTP, 0.2 mM dCTP or dmCTP, 0.2 mM dTTP, with 50 μ g ml⁻¹ rifampicin present to prevent re-formation of ternary complexes by way of newly initiated transcription, and 10 units of *HaeIII* added to limit the extent of DNA synthesis by linearizing the DNA template. KMnO_4 was then added to a final concentration of 5.1 mM. After 1 min at 37 °C, the KMnO_4 reaction was quenched with 5 μ l 14 M β -mercaptoethanol. The sample was treated with 80 μ g ml⁻¹ proteinase K in the presence of 0.5% SDS for 30 min at 37 °C, followed by phenol-chloroform extraction and ethanol precipitation with 15 μ g ml⁻¹ glycogen as carrier. The pellet was dissolved in a buffer (10 mM Bis-Tris-propane-HCl, 10 mM MgCl_2 , 1 mM DTT, pH 7.0) that allowed optimal digestion by *AluI* (10 units) during 8 min at 37 °C. Primer extension was done with a 5' end-labelled 19-nt single-stranded DNA complementary to the non-template strand; subsequent sample preparation and electrophoresis on a 10% polyacrylamide gel (37.5: 1 acrylamide: bisacrylamide) with 8 M urea in 1 $\times$ TBE were performed as described⁴²⁻⁴⁴.



one round of replication has occurred. Thus, the replication fork is indeed able to pass the ternary complex without displacing it.

When we briefly added rNTPs to mixtures that had finished replication, nascent RNA was detected on many extensively replicated DNA templates (Fig. 5*h*), indicating that by-passed ternary complexes are functional. Finally, DNA structures other than those expected from rolling-circle replication were not observed, arguing against the possibility that any of the findings in this article are explained by some unanticipated replication mechanism.

Discussion

Our examination of the consequences of a collision between a replication fork and a codirectionally orientated, stalled RNA polymerase ternary transcription complex yields a surprising result: the replication fork passes the ternary complex after only a brief pause (estimated to last <1 s; Fig. 1*b*); the by-passed ternary complex not only remains bound to the DNA (Fig. 2)

with a transcription bubble at its original DNA site (Fig. 3), but it is fully competent to resume RNA synthesis (Fig. 4). Electron microscope examination of the reaction products supports this conclusion at the single macromolecule level (Fig. 5).

Our results do not merely reflect a special property of T4 late gene transcription. We have repeated the experiment shown in Fig. 4 with a ternary transcription complex derived from initiation at a σ^{70} promoter by *E. coli* RNA polymerase (in the absence of any T4 protein) and have obtained the same result (data not shown).

A stalled ternary complex is an imperfect representation of true transcription intermediates, whose normal structures are likely to be kinetically determined⁴. But the recovery of full-length transcripts during a chase in which rolling circle replication is ongoing (Fig. 4*a*) suggests that many transcription intermediates (and not just our stalled ternary complex) can survive the replication fork; further evidence supporting this point will be presented elsewhere.

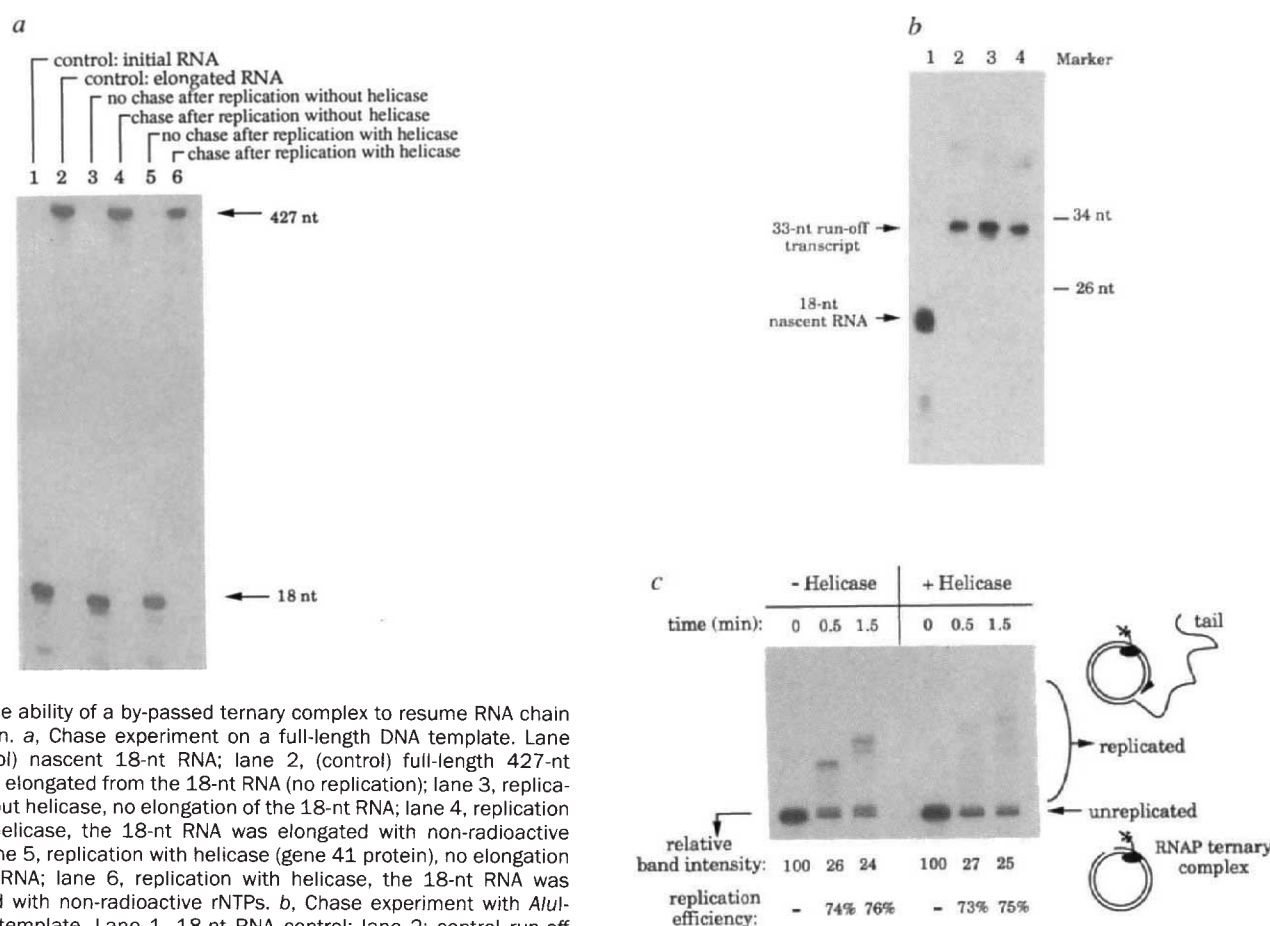


FIG. 4 The ability of a by-passed ternary complex to resume RNA chain elongation. *a*, Chase experiment on a full-length DNA template. Lane 1, (control) nascent 18-nt RNA; lane 2, (control) full-length 427-nt transcript elongated from the 18-nt RNA (no replication); lane 3, replication without helicase, no elongation of the 18-nt RNA; lane 4, replication without helicase, the 18-nt RNA was elongated with non-radioactive rNTPs; lane 5, replication with helicase (gene 41 protein), no elongation of 18-nt RNA; lane 6, replication with helicase, the 18-nt RNA was elongated with non-radioactive rNTPs. *b*, Chase experiment with *Alul*-cut DNA template. Lane 1, 18-nt RNA control; lane 2; control run-off transcript (~33-nt RNA); lane 3, run-off transcript after replication without helicase; lane 4, run-off transcript after replication with helicase. *c*, Determination of replication efficiency. Replication (with or without DNA helicase) proceeded at 37 °C for the time indicated. The replication efficiencies are calculated from the reduction of the radioactive signal (quantified using a PhosphorImager) at the position of the non-replicated molecules.

METHODS. *a*, *In vitro* replication was as described in Fig. 2 (except that the DNA was not linearized) for 5 min (without helicases) or for 2 min (with helicases), followed by the addition of cold rNTPs (0.5 mM rATP, 0.5 mM rGTP, 0.2 mM rCTP and 0.2 mM rUTP) to chase the nascent transcript at 37 °C for 8 min. Samples were then chilled on ice, treated with 2 units of DNase I (with CaCl₂ at a final concentration of 0.5 mM), phenol-chloroform extracted and electrophoresed on a 10% denaturing polyacrylamide gel. To rule out the possibility of RNA polymerase reassociation during the rNTP chase, this experiment has been repeated:

(1) in the presence of rifampicin (30–50 $\mu\text{g ml}^{-1}$); (2) with synthetic oligonucleotides containing the promoter sequences in 10–20-fold molar excess over the template; (3) with yeast ribosomal RNA (40–100 mg ml^{-1}). These variations did not change the outcome of the experiment. *b*, As in *a*, except that RNA chains were elongated on templates that had been digested with 10 units of *Alul*. *c*, *In vitro* replication was done on the column-purified complex under the same conditions of the chase experiments described in *a* and *b*. The reaction was stopped by chilling the samples on ice at the indicated times. Heparin and Ficoll were added to concentrations of 100 $\mu\text{g ml}^{-1}$ and 3%, respectively. Samples were then loaded on a 0.8% neutral agarose gel (non-denaturing) in 1 $\times$ TBE and electrophoresed at room temperature for 4–5 h at 7 V cm^{-1} . The gel was dried and autoradiographed or exposed to a PhosphorImager screen for quantitative analysis.

It is not obvious how two bulky enzyme complexes can pass one another in a non-destructive way. In particular, because the codirectionally moving DNA and RNA polymerases use the same DNA single strand as template, the replication apparatus almost certainly unwinds the end of the growing RNA transcript that is base-paired to DNA. How difficult is this likely to be? Recent structural studies on transcription complexes^{17,18} suggest that growing RNA chains may be relatively loosely held at the 3' end in a short DNA-RNA hybrid, and periodically transferred to a tighter binding site in *E. coli* RNA polymerase¹⁷⁻²⁵. This would imply that: (1) the contribution of the RNA-DNA hybrid to the stability of the ternary complex need not be as great as previously thought^{26,27}, making its transient unpairing less daunting; (2) the RNA-DNA hybrid need not present an insurmountable barrier to the progression of a replication fork.

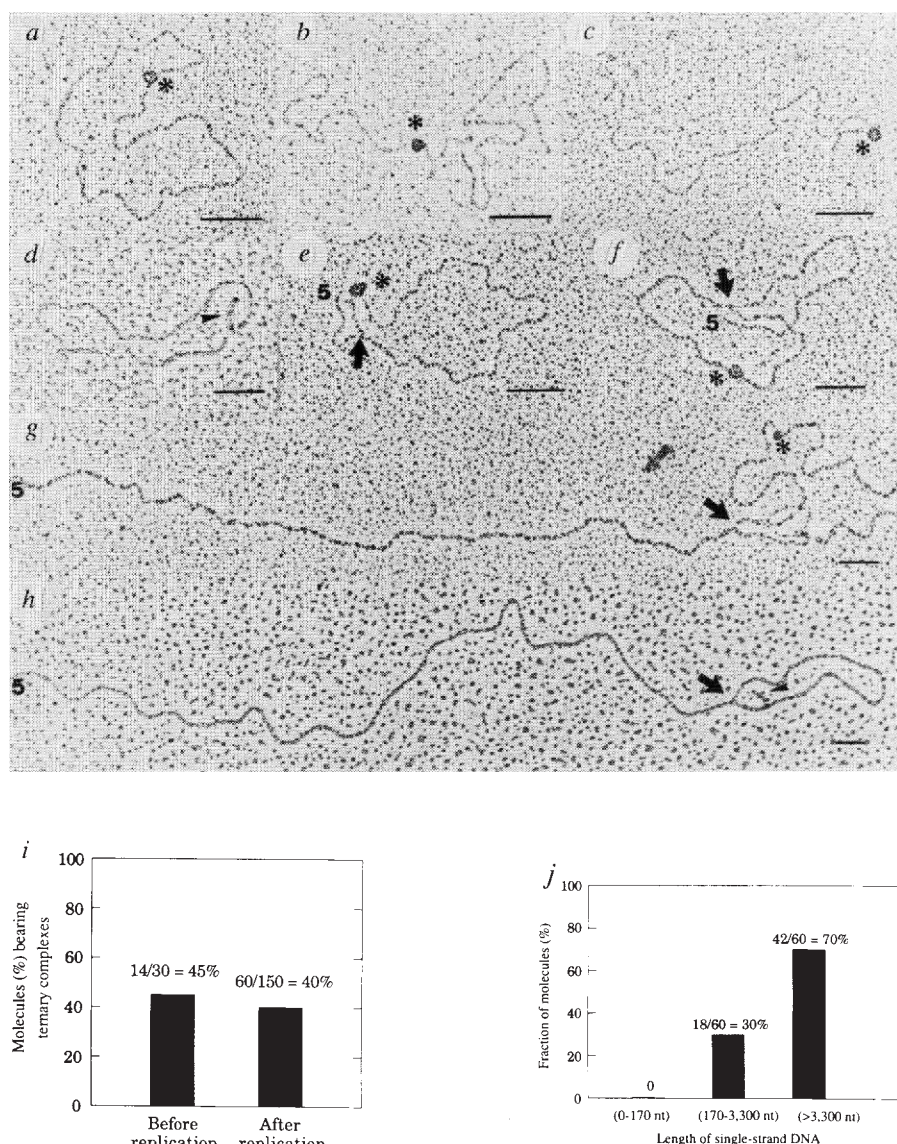
Although we cannot rule out the possibility that some property of the T4 DNA replication proteins is important for our observations (for example, tethering the RNA polymerase to the DNA molecule, thereby preventing its release as the DNA

polymerase passes), it seems more likely that our results are intrinsic to the *E. coli* RNA polymerase, a multiple-subunit enzyme²⁸ that can undergo large conformational changes²⁹. A schematic model of this type is shown in Fig. 6, where at least two DNA-interacting domains of the polymerase are present within the ternary complex, each individually detachable from the DNA without destroying the complex. When the replication fork invades the interior of the complex, it causes a momentary unpairing of the short RNA-DNA hybrid at the 3' end of the nascent RNA in the ternary complex, but the other DNA-binding domain keeps the RNA polymerase attached to the daughter DNA helix. The nascent RNA remains bound to RNA polymerase, and it may ensure the maintenance of transcription fidelity by its specific hybridization back to the DNA template.

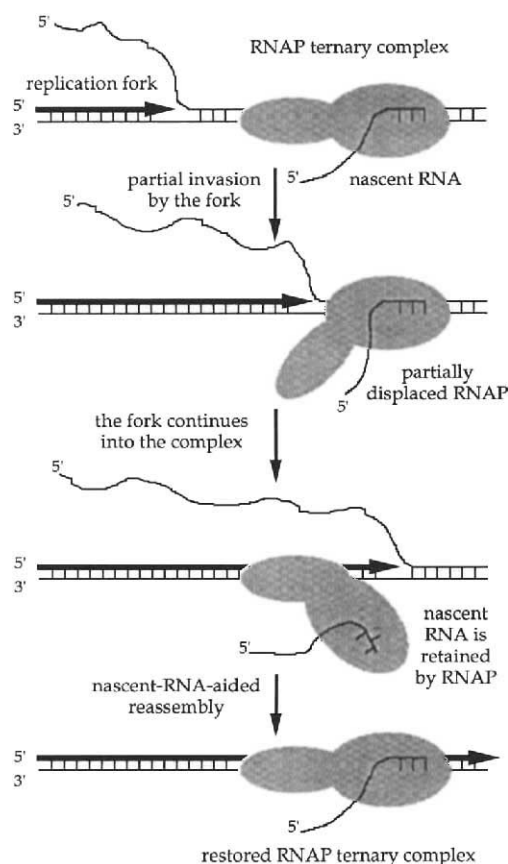
In contrast to the ternary transcription complex, a promoter-bound RNA polymerase that is not transcribing is displaced from the template after replication³⁰. Compared to the very tight ternary complex that enables RNA polymerase to transcribe in a highly processive manner, promoter binding by RNA polymer-

FIG. 5 Electron microscopic examination of replication products. *a*, A non-replicated molecule bearing a globular particle (the putative RNA polymerase ternary complex). *b* and *c*, Mapping the position of the globular particle by restriction enzyme digestion (*b*, *SspI*; *c*, *HindIII*). The relative distances from the particle to the two ends (long/short): for *SspI*, the measured ratio is 1.8 (± 0.1) (expected 1.8); for *HindIII*, the measured ratio is 23 (± 3) (expected 25). *d*, Production of nascent RNA. *e-g*, Various extents of replication take place on templates bearing a ternary complex. The ternary complex remains on the replicated DNA. *h*, Production of RNA on the extensively replicated DNA template. *i*, Comparison of the percentage of molecules bearing a ternary complex before and after replication. Randomly selected samples of 30 non-replicated and 150 replicated molecules were examined. *j*, Distribution of replicated molecules with tails of varying lengths. The distance from the gene 2 nick (replication origin) to the ternary complex is ~ 170 nt. If the replication fork stalls permanently before the ternary complex, no molecules should bear a tail exceeding this length. Scale bar, 0.1 μm . Arrows, replication forks. Asterisks, RNA polymerase (*a-c*, *e-g*). Arrowheads, nascent RNA (*d* and *h*). '5', the 5' end of a displaced DNA tail.

METHODS. *In vitro* replication was done on CL-2B-column-purified ternary complex as described in the legend to Fig. 1*b*. The reaction was stopped by chilling the sample on ice, followed by the addition of NaCl to 0.5 M and passage through CL-2B to remove replication proteins. Electron microscopy studies were done on the radioactive fractions as follows: *a-c*, and *e-g*: 2–8 μl samples were applied to glow-discharged carbon grids for 2 min, dehydrated in 100% ethanol and uranyl acetate, and shadowed with platinum at an angle of 8 degrees⁴⁵; *d* and *h*, rNTPs (1 μM each) were added to elongate the nascent 18-nt RNA at 37 °C for 30 s. The reaction was stopped with 20 mM Na₃-EDTA. Transcripts were crosslinked to DNA templates



by ultraviolet light (254 nm) irradiation at 25 °C for 10 min at a distance of ~ 2 cm from a UVGL-25 lamp. Samples were spread with cytochrome *c* as described in ref. 46 before examining them with a Philips EM400 microscope.



ase is a weaker interaction. Because it relies on hydrogen-bonding interactions with specific bases on both DNA strands⁴, the separation of the two strands of the double helix during replication would be expected to destabilize the promoter complex. There is little energy investment during promoter binding by RNA polymerase, and its displacement is less costly to the cell.

Our results suggest the existence of a mechanism that allows the RNA polymerase and its attached transcript to survive the collision between the replication and transcription machineries.

FIG. 6 A schematic model to account for some of the experimental observations. Well-separated DNA-binding domains might allow the *E. coli* RNA polymerase (RNAP) to retain its place as replication passes through. Replication proteins and the transcription bubble are not drawn. As described in the text, the retention of exact transcription register that we observed (Fig. 4b) is likely to involve base-pair reformation by the 3' end of nascent RNA. In addition, the ternary complex contains about 17 base pairs of separated DNA strands⁴⁷ in the form of a transcription bubble. Because there is a substantial energetic cost to reforming this bubble⁴, the passage of a replication fork through the ternary transcription complex might involve a reaction pathway that never entirely dissipates the DNA strand separation.

Because the *E. coli* and the eukaryotic nuclear RNA polymerases have evolved from a common ancestor and have homologous subunits that share amino-acid sequence homology^{31,32}, our results may also be relevant to the behaviour of these polymerases when a replication fork passes.

Our conclusions are entirely based on *in vitro* experiments done with highly purified proteins. Is there any *in vivo* evidence for or against a mechanism of this type? When the fate of the nascent transcript of a large *Drosophila* gene (*Ubx*) whose complete transcription takes longer than the time of a cell cycle was observed³³, DNA synthesis *in vivo* did not abort the nascent transcript (in this case, the orientation of the fork relative to the RNA polymerase movement is unknown). In contrast, electron microscopy has been used to show that the nascent transcripts of a rRNA gene of *E. coli* are displaced from the template when a replication fork invades the transcription unit from either direction³⁴. But these rRNA transcripts are unusual in at least two aspects: they are attached to closely spaced RNA polymerase molecules, and they are modified by a set of specialized RNA-binding proteins³⁵. We would predict that a different result would be obtained with other transcription units. Methods that permit a quantitative analysis, such as simultaneously probing the fork movement and nascent transcript production by nucleic acid hybridization, should be useful for examining this issue.

It is certainly possible that unidentified protein factors exist *in vivo* that modulate the basic mode of interaction between a replication fork and RNA polymerase observed in our experiments. Nevertheless, the fact that a ternary complex can survive a replication fork *in vitro* demonstrates a remarkable ability of RNA polymerase to cope with perturbing events during its elongation of RNA chains. □

Received 16 June; accepted 30 September 1993.

- Bremer, H. & Dennis, P. P. in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (eds Neidhardt, F. C.) 1527–1542 (Am. Soc. of Microbiology, Washington DC, 1987).
- Hirose, S., Hiraga, S. & Okazaki, T. *Molec. gen. Genet.* **189**, 422–431 (1983).
- Johnston, D. E. & McClure, W. R. in *RNA Polymerase* (eds Losick, R. & Chamberlin, M. J.) 413–428 (Cold Spring Harbor Laboratory Press, New York, 1976).
- von Hippel, P. H. et al. *Rev. Biochem.* **53**, 389–446 (1984).
- Wang, W., Carey, M. & Gralla, J. D. *Science* **255**, 450–453 (1992).
- Christensen, A. C. & Young, E. T. in *The Bacteriophage T4* (eds Matthews, C. K. et al.) 184–188 (Am. Soc. for Microbiology, Washington DC, 1983).
- Kassavetis, G. A. & Geiduschek, E. P. *EMBO J.* **1**, 107–114 (1982).
- Kassavetis, G. A., Zenter, P. G. & Geiduschek, E. P. *J. biol. Chem.* **261**, 14256–14265 (1986).
- Kassavetis, G. A. & Geiduschek, E. P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5101–5105 (1984).
- Hinkle, D. C., Ring, J. & Chamberlin, M. J. *J. molec. Biol.* **70**, 197–207 (1972).
- Williams, K. P., Kassavetis, G. A. & Geiduschek, E. P. *J. biol. Chem.* **262**, 12365–12371 (1987).
- Alberts, B. M. *Phil. Trans. R. Soc.* **B317**, 395–420 (1987).
- Alberts, B. M. & Frey, L. *Nature* **227**, 1313–1318 (1970).
- Morris, C. F., Hamma-Inaba, H., Mace, D., Sinha, N. K. & Alberts, B. M. *J. biol. Chem.* **254**, 6787–6796 (1979).
- Morris, C. F., Moran, L. A. & Alberts, B. M. *J. biol. Chem.* **254**, 6796–6802 (1987).
- Kassavetis, G. A., Blanco, J. A., Johnson, T. E. & Geiduschek, E. P. *J. molec. Biol.* **226**, 47–58 (1992).
- Rice, G. A., Kane, C. M. & Chamberlin, M. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4245–4249 (1991).
- Surratt, C. K., Milan, S. C. & Chamberlin, M. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7983–7987 (1991).
- Kumar, S. A. & Krakow, J. S. *J. biol. Chem.* **250**, 2879–2884 (1975).
- Kumar, S. A. *Prog. Biophys. molec. Biol.* **38**, 165–210 (1981).
- Bean, B., Koren, R. & Milder, A. S. *Biochemistry* **16**, 3322–3333 (1977).
- Stein, B. J. & Milder, A. S. *Biochemistry* **17**, 2675–2684 (1978).
- Borukhov, S., Sagitov, V. & Goldfarb, A. *Cell* **72**, 459–466 (1993).

- Fisher, R. & Blumenthal, T. *J. biol. Chem.* **257**, 1702–1704 (1982).
- Mi, H. & Hartmann, G. R. *Eur. J. Biochem.* **131**, 113–118 (1983).
- von Hippel, P. H. & Yager, Y. D. *Science* **255**, 809–812 (1992).
- Yager, T. D. & von Hippel, P. H. *Biochemistry* **30**, 1097–1118 (1991).
- Chamberlin, M. J. in *RNA Polymerase* (eds Losick, R. & Chamberlin, M. J.) 17–67 (Cold Spring Harbor Laboratory Press, New York, 1976).
- Joe, J.-H., Burgess, R. R. & Record, M. T. *J. molec. Biol.* **184**, 441–453 (1985).
- Bedinger, P., Hochstrasser, M., Jongeneel, V. G. & Alberts, B. M. *Cell* **34**, 115–123 (1987).
- Sentenac, A. et al. in *Transcription Regulation* (eds McKnight, S. L. & Yamamoto, K. R.) 27–54 (Cold Spring Harbor Laboratory Press, New York, 1992).
- Young, R. A. *Rev. Biochem.* **60**, 689–715 (1991).
- Shermoen, A. W. & O'Farrell, P. H. *Cell* **67**, 303–310 (1991).
- French, S. *Science* **258**, 1362–1365 (1992).
- Richardson, J. P. *Crit. Rev. Biochem. molec. Biol.* **28**, 1–30 (1993).
- Nossal, N. G. *J. biol. Chem.* **255**, 2176–2182 (1980).
- Cha, T.-A. & Alberts, B. M. *J. biol. Chem.* **261**, 7001–7010 (1986).
- Herendeen, D. R., Kassavetis, G. A., Barry, J., Alberts, B. M. & Geiduschek, E. P. *Science* **245**, 952–958 (1989).
- Kunkel, T. A., Roberts, J. D. & Zakour, R. A. *Meth. Enzym.* **154**, 367–382 (1987).
- Meyer, T. F., Geider, K., Kurz, C. & Schaller, H. *Nature* **278**, 365–367 (1979).
- Hayatsu, H. & Ukita, T. *Biochem. biophys. Res. Commun.* **29**, 556–561 (1967).
- Gralla, J. D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3078–3081 (1985).
- Sasse-Dwight, S. & Gralla, J. D. *J. molec. Biol.* **202**, 107–119 (1988).
- Ide, H., Kow, Y. W. & Wallace, S. S. *Nucleic Acids Res.* **13**, 8035–8052 (1985).
- Dubochet, J., Ducommun, M., Zollinger, M. & Kellenberger, E. *Ultrastruct. Res.* **35**, 147–167 (1971).
- Kleinschmidt, A. K. *Meth. Enzym.* **12**, 361–377 (1968).
- Gamper, H. B. & Hearst, J. E. *Cell* **29**, 81–90 (1982).

ACKNOWLEDGEMENTS. We thank J. Barry and G. A. Kassavetis for advice and help, and K. Hacker, G. Sanders, J. Raff, C. Mirzayan, R. Schrock, R. V. Arlian and J. Cox for reading the manuscript. This work is supported by separate grants from the National Institute of General Medical Sciences to B.M.A. and E.P.G., and by the Howard Hughes Medical Institute (M.L.W.). B.M.A. is an American Cancer Society Research Professor.

The intrinsic luminosity of γ -ray bursts and their host galaxies

E. E. Fenimore*, R. I. Epstein*, C. Ho*,
R. W. Klebesadel*, C. Lacey*, J. G. Laros*,
M. Meier*, T. Strohmayer*, G. Pendleton†,
G. Fishman†, C. Kouveliotou† & C. Meegan†

* Los Alamos National Laboratory, Los Alamos,
New Mexico 87545, USA

† Marshall Space Flight Center, Huntsville, Alabama 35812, USA

THE Burst and Transient Source Experiment (BATSE) on the Compton Gamma-Ray Observatory has shown that, although γ -ray bursts are distributed isotropically on the sky, there is an apparent dearth of weak events compared to those expected from a homogeneous distribution of sources^{1,2}. This suggests that the bursts originate either in an extended galactic halo³ (that is, locally) or at cosmological distances^{4,5}. The intensity distribution of the bursts can be used to constrain their source properties and spatial distribution, and here we address this question by considering the BATSE data with those from the Pioneer Venus Orbiter, correcting for the different response of the two instruments. We show that the composite intensity distribution can after all be fitted by a simple, homogeneous model of γ -ray burst sources with identical intrinsic luminosities where the faintest BATSE events are at a redshift of $z \sim 0.80 \pm 0.05$. We obtain an upper limit of about -18 on the absolute magnitude of the host galaxies by assigning model-derived distances to the brightest extragalactic objects found¹⁵ within the error boxes of eight well localized γ -ray bursts. All but the faintest active galaxies are excluded as the source of the bursts. The bursts may instead be associated with normal galaxies, but only if the host-galaxy magnitudes are close to our upper limit.

If γ -ray bursters (GRBs) are at cosmological distances, the dearth of weak events results from the expansion of the universe, either because the signal is redshifted out of the bandpass of the instrument, and/or through evolution of the sources. The relative number of bursts at different intensities can be used to identify which cosmological models are acceptable, assuming the same luminosity for all sources. Studies have shown that the BATSE distribution by itself is consistent with a wide range of cosmologies^{6,7} and evolutionary models⁸. Comparisons^{9,10} of the Pioneer Venus Orbiter (PVO) and the BATSE threshold burst rates gave estimates of the intrinsic luminosity that disagreed by a factor of 10. Particular cosmological models can be compared to the observations with a log N -log P distribution. In log N -log P distributions, N is the integral number of events observed above some peak flux P (either P_e for $\text{erg cm}^{-2} \text{s}^{-1}$, or P_p for $\text{photon cm}^{-2} \text{s}^{-1}$). Distributions of log N -log P have the disadvantage that the number of weak events for each instrument could be underestimated because the threshold depends on the time varying background. By combining the log N -log P distributions from the PVO and BATSE, correcting for the fraction of time the instrument is capable of detecting bursts (that is, the livetime) and threshold effects, we obtain a single distribution that covers a dynamic range of 10^3 . The PVO observed for ~ 14 years with a 90% livetime¹¹ and a 4π field of view whereas BATSE sees the 2.6π portion of the sky not blocked by the Earth and has a livetime of 55% (ref. 12). It will take BATSE 30 more years to achieve the dynamic range of the combined PVO/BATSE distribution. Although early attempts⁹ to compare the BATSE and PVO data implied that the BATSE burst detection rate was too low (at a 3σ level), our re-examination of the effective lifetime of PVO brought the PVO and BATSE results closer together¹¹.

To combine the BATSE and PVO databases into a single log

N -log P distribution, one must select the events uniformly and calculate the fluxes in a similar manner. The two instruments differ significantly in how the events are selected and the selection criteria we have used are described in the box.

Figure 1a shows the resulting log N -log P_p histograms for the 86 BATSE events and 146 PVO events selected. Without any arbitrary scaling, there is excellent agreement between the two curves. Six events seen by both PVO and BATSE had peak fluxes that agreed within the statistics. Unfortunately the events are so weak in PVO that these events cannot be used to independently determine the relative normalization of the BATSE and PVO distributions.

We have fit the combined data to models^{9,13} that assume an $\Omega_0 = 1$ cosmology, a constant number of events per unit volume-time, and a constant intrinsic peak luminosity (that is, a standard candle). The dotted and dashed vertical lines show the intensity intervals ($P_{p,i}, P_{p,i+1}$) for which the differential number of observed events are compared to the model for BATSE and PVO, respectively. The number of expected events between the (photon $\text{cm}^{-2} \text{s}^{-1}$) values of $P_{p,i}$ and $P_{p,i+1}$ is

$$\Delta N(P_{p,i}, P_{p,i+1}) = \int_{r_i}^{r_{i+1}} \left[\frac{2}{2 - rH_0/c} \right]^{-2} 4\pi\rho r^2 dr \quad (1)$$

where ρ is the rate-density of events per comoving volume (bursts $\text{yr}^{-1} \text{Gpc}^{-3}$) and r_i is the distance to the bursts with

BOX 1 BATSE and PVO event selection

THE PVO triggers in the energy range of 100–2,000 keV on timescales of 1/4, 1 and 4 s and most events have their peak intensities known at 0.1875-s intervals. BATSE triggers in the energy range 50–300 keV on timescales of 64, 256 and 1,024 ms and most events have their peak intensity known for sliding time periods of 64, 256, 1,024 ms. In addition, a burst brighter than an event in BATSE's memory can overwrite the memory, thus increasing the effective threshold for the brighter event above that expected from the background alone.

Because both instruments trigger on 1/4 and 1 s timescales, we excluded events from BATSE that did not have either the 1/4 s or the 1 s $C_{\text{max}}/C_{\text{min}}$ (peak count rate divided by detection threshold) greater than unity. We also exclude those events that overwrote other events. The data required to obtain the peak fluxes are not overwritten so those events were not excluded. These exclusions reduce the number of usable BATSE events between 5 March and 21 April 1992 from 260 to 197. The PVO instrument does not report the timescale on which it triggered; a re-analysis has estimated $C_{\text{max}}/C_{\text{min}}$ for each timescale and events are excluded when neither the 1/4 s nor the 1 s $C_{\text{max}}/C_{\text{min}}$ is greater than unity. We include only PVO events that occurred during the period from 14 September 1978 to 1 January 1992, which reduces the number of events from 295 to 265. BATSE typically reported peak fluxes in the bandpass 30–500 keV and the PVO peak fluxes are in the bandpass 100–2,000 keV. For the purposes of combining the two databases, we used the peak fluxes in the range 100–500 keV. The BATSE peak fluxes were determined for a sliding 256-ms interval and the PVO fluxes were determined at 250 ms intervals. Despite our efforts to cross-calibrate the two instruments, a $\pm 10\%$ systematic uncertainty in the relative normalization is likely.

The low-flux end of log N -log P distributions can be distorted by instrumental effects. To avoid threshold effects, we have established lower limits for including events in the analysis above which threshold effects are insignificant. For BATSE the limits are 1.0 photon $\text{cm}^{-2} \text{s}^{-1}$ and $2.8 \times 10^{-7} \text{ erg cm}^{-2} \text{s}^{-1}$ which result in 86 and 99 events, respectively. For PVO, the limits are 20 photon $\text{cm}^{-2} \text{s}^{-1}$ and $8 \times 10^{-6} \text{ erg cm}^{-2} \text{s}^{-1}$ which results in 146 and 144 events, respectively. The normalization of the BATSE curve assumes a solid angle coverage of 2.6π , 318 days of observing and a livetime fraction for triggered, non-excluded bursts equal to 0.55×0.76 , where the 0.76 factor accounts for the exclusion of 63 of the 260 events. The normalization of the PVO curve assumes a solid angle coverage of 4π , 4,840 days of observing and a livetime fraction for triggered, non-excluded bursts equal to 0.75.

TABLE 1 Limits on GRBs and their hosts

GRB	erg cm ⁻² s ⁻¹	Log N-log P distance (Gpc)	Schaefer* distance (Gpc)	M _v	Fraction of stellar mass
781119	2.3 × 10 ⁻⁵	0.33	3.2†	-16.9	0.23
781124	1.7 × 10 ⁻⁵	0.37	2.5	-17.6	0.42
790325	7.8 × 10 ⁻⁶	0.53	2.5‡	-18.4	0.29
790329	4.6 × 10 ⁻⁶	0.66	1.0	-20.9	0.95
790406	9.9 × 10 ⁻⁶	0.48	4.5	-16.9	0.23
790418	1.6 × 10 ⁻⁵	0.39	2.1	-18.1	0.34
790613	3.8 × 10 ⁻⁶	0.73	1.9	-19.7	0.84
791116	2.4 × 10 ⁻⁵	0.32	3.4	-16.7	0.25

A galaxy mass function is constructed from the luminosity function¹⁷ by assuming that stellar mass scales with luminosity. The fraction of stellar mass (column 6) is the ratio of the integral of the mass function above M_v (column 5) to the total area under the curve.

* From Schaefer¹⁵ assuming M31 luminosity. † The distance was based on a galactic star; 3.2 Gpc is the distance¹⁶ to the brightest extragalactic object in the error box, a QSO (B. E. Schaefer, personal communication). ‡ 2.5 Gpc is for the brightest extragalactic object¹⁵.

intensity $P_{p,i}$ found from

$$P_{p,i} = \frac{L_{50} 10^{50}}{\int_{30}^{2,000} E \phi(E) dE} \frac{\int_{E_1}^{E_2} \phi \left(\left[\frac{2}{2 - r_i H_0 / c} \right]^2 E \right) dE}{4\pi r_i^2} \quad (2)$$

Here, $\phi(E)$ is the assumed spectral shape at the source (for which we have used a thermal bremsstrahlung spectrum, $E^{-1} e^{-E/kT}$, with $kT = 350$ keV), E_1 to E_2 is our bandpass (100–500 keV), H_0 is the Hubble constant (~ 75 km s⁻¹ Mpc⁻³), c is the speed of light, and L_{50} is the peak luminosity at the source in units of 10^{50} erg s⁻¹ in the bandpass 30–2,000 keV. In these models there are two free parameters (ρ, L_{50}). The best fit cosmological model is shown in Fig. 1 as a solid curve and is consistent with both the BATSE and PVO data ($\chi^2 = 9.0$ with nine degrees of freedom, dof). The dotted curve in the insert in Fig. 1a shows the 1 σ confidence region for L_{50} and $\rho L_{50}^{3/2}$ (in units of bursts yr⁻¹ Gpc⁻³). The best fit $L_0 = L_{50} 10^{50}$ is 5.2×10^{50} erg s⁻¹ h_{75}^{-2} and $\rho = 24$ bursts yr⁻¹ Gpc⁻³ h_{75}^3 where $h_{75} = H_0/75$. The 1 σ confidence region has a χ^2 that is 2.3 larger than the χ^2 at the minimum¹⁴. The maximum extent of L_{50} in the 1 σ region is $5.2^{+2.3}_{-1.9}$ which are conservative limits on L_{50} (see ref. 14). A range for the single parameter L_{50} without restricting ρ which will contain the true model 68% of the time is found from where χ^2 increases by 1; this gives $5.2^{+1.4}_{-1.3}$. Preliminary results using an additional 56 BATSE events in 4 bins between 0.5 and 1.0 photon cm⁻² s⁻¹ and an estimate of the BATSE threshold efficiency changes L_{50} to 4.6, within the uncertainty of the fit found in Fig. 1a. The 10% uncertainty in the normalization of the two instruments can cause a ± 1 variation in L_{50} .

A successful model needs to be consistent with both log N -log P_p and log N -log P_e . The P_p and P_e fluxes are not completely independent since they both arise from the same data. However, additional information (the spectral shape) is used to determine P_e and the lack of independence should not cause a large effect on the size of the joint confidence region. The log N -log P_e distribution is found from equations (1) and (2) except that there is an additional factor of E in the numerator's integral of equation (2). Figure 1b shows the best fit model, $L_0 = 7.2^{+1.6}_{-1.5} \times 10^{50}$ erg s⁻¹ h_{75}^{-2} and $\rho = 20$ bursts yr⁻¹ Gpc⁻³ h_{75}^3 . The fit is acceptable ($\chi^2 = 8.5$ with 11 dof defined above) and the dashed curve in the insert of Fig. 1b shows the 1 σ confidence region. The model that best fits both distributions has $L_0 = 6.0^{+1.1}_{-0.8} \times 10^{50}$ erg s⁻¹ h_{75}^{-2} and $\rho = 22$ bursts yr⁻¹ Gpc⁻³ h_{75}^3 with an acceptable χ^2 (20.9 with 22 dof). The 1 σ confidence region of the joint fit is shown in both inserts as a solid line. This best fit implies that the magnitude of the cosmological redshift at the threshold of BATSE (0.7 photon cm⁻² s⁻¹) is $z = 0.79 \pm 0.05$

and at the threshold of PVO (10 photon cm⁻² s⁻¹) is $z = 0.22 \pm 0.02$.

Table 1 lists eight GRBs with well defined locations from the interplanetary network for which there have been deep searches for counterparts^{15,16}. Based on the P_e used in Fig. 1b for these events (column 2) and $L_0 = 6 \times 10^{50}$ erg s⁻¹, the distance consistent with Fig. 1 is estimated from equation (2) for each event (see column 3). Visual apparent magnitudes are not available for all of the brightest object in the error boxes. Rather, apparent magnitudes for a variety of bandpasses have been reported by Schaefer and used to estimate a lower limit on the distance to the object assuming an M31 luminosity and spectrum¹⁵ (see column 4). We have used our distance (column 3), the absolute visual magnitude of M31, and the Schaefer distance to estimate the absolute visual magnitude of the brightest object within the error box (see column 5). The average absolute magnitude is $-18 + 5 \log(h_{75})$.

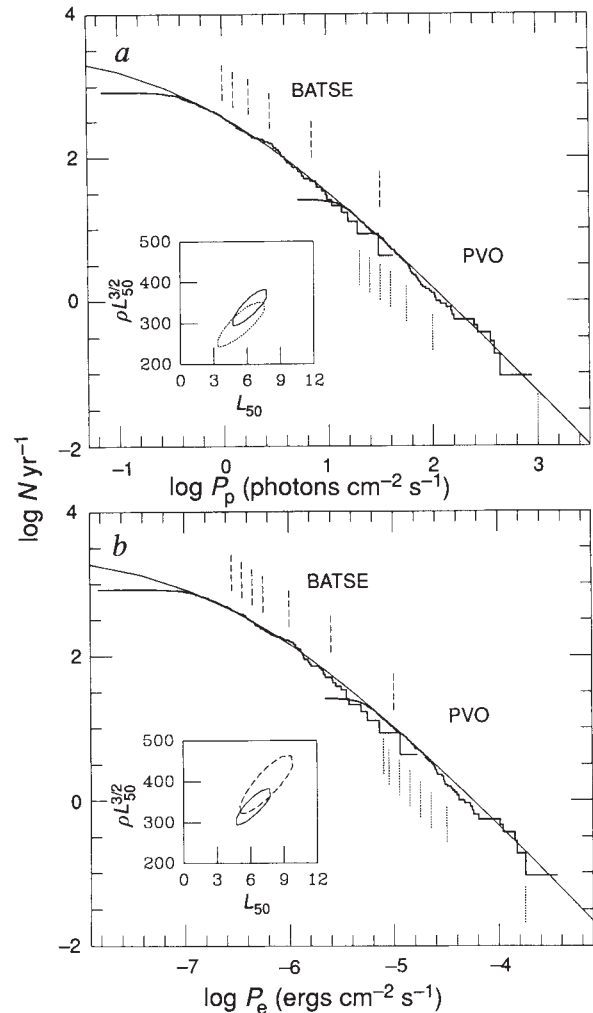


FIG. 1 Distribution of log N -log P_p for uniformly selected events from the PVO and BATSE. a, The two histograms are the log N -log P_p distributions. The dashed and dotted vertical lines are for the BATSE and PVO, respectively, and indicate the ranges over which events were binned to compare to models. The log N -log P_p distributions below the lowest dotted and dashed lines deviate from the model due to threshold effects and were not used in the fit. The solid line shows an acceptable standard candle cosmological model that fits both sets of data. The dashed line in the inset shows the 1 σ confidence region for the model. b, Similar to a except for the erg cm⁻² s⁻¹ distributions. The dotted region in the insert is the confidence region for the erg cm⁻² s⁻¹ fit. In both inserts is shown a solid curve which is the 1 σ confidence region for a simultaneous fit to both the erg cm⁻² s⁻¹ and photon cm⁻² s⁻¹ distributions. The resulting best fit standard candle luminosity is quite small ($6^{+1.1}_{-0.8} \times 10^{50}$ erg s⁻¹).

These absolute magnitudes are surprisingly small if GRBs are located in regions with a high density of stars (for example, active galactic nuclei). They are even small compared to what one might expect from GRBs being associated with normal galaxies. Assuming that the number of stars in a galaxy scales with its luminosity, and using an observed luminosity function for galaxies¹⁷, we have used a test similar to $V/V_{\max}$ (see ref. 18) to check the consistency of the absolute visual magnitudes with those of normal galaxies. In column 6 (Table 1) we show the fraction of stellar mass that occurs in galaxies with magnitudes larger than shown in column 5; this fraction is reported in column 6. If GRBs are associated with a population whose density scales with normal stars, column 6 would be distributed uniformly between 0 and 1. The average of the fractions is 0.5, as expected for a uniform distribution. Most galaxies with an absolute magnitude of ~ -18 are spirals¹⁷ whereas all of the objects reported by Schaefer were point-like and inconsistent with being spirals. Thus, the absolute magnitudes listed in column 5 must be considered upper limits. Using the statistics of uniform distributions, if the actual hosts are more than 1 magnitude fainter (on average), then the observed distribution would be inconsistent with GRBs coming from normal galaxies ($<10\%$ chance than the observed upper limits would occur by random chance). The statistical and systematic uncertainties on L_0 could allow the actual hosts to be 1.5 magnitude fainter rather than 1, but that would still mean that the parameter space where normal galaxies are consistent with the observations is small. If re-evaluations of the error boxes listed in Table 1 fail to show extragalactic objects within 1.5 magnitudes of column 5, then one can reject the hypothesis that GRBs are associated with objects (such as galaxies) whose density scales with normal matter. In that case, either GRBs are orphans (not associated with normal matter) or our estimate of L_0 is too small.

Based on 22 GRBs $\text{Gpc}^{-3} \text{yr}^{-1}$, we estimate that the bursts occur at a rate of $\sim 10^{-20}$ GRBs $M_{\odot}^{-1} \text{yr}^{-1}$.

The log N -log P distribution can be modified by either number evolution (that is, $\rho = \rho(z)$) or luminosity evolution (that is, $L_{50} = L_{50}(z)$). (Spectral (that is, kT) evolution has little effect on L_{50} .) Perhaps $\rho(z)$ and $L_{50}(z)$ conspire to produce the appearance of a log N -log P with a small L_{50} where, in fact, L_{50} is larger, allowing a brighter absolute magnitude for the host galaxy. To investigate this possibility, we have assumed that ρ and L_{50} scale as some power of $1+z$ (let p_{ρ} and p_L represent the respective exponents). With pure luminosity evolution, our model accommodates p_L as large as 1.9 (rejection at 90% confidence), with a resulting maximum best-fit L_{50} at $z=0$ of 13.6 (a decrease in the absolute magnitude of the hosts of 0.9). For pure density evolution we find that all models with $p_{\rho} \geq 3.9$ are rejected at least at the 90% confidence level. This gives a maximum value of $L_{50} = 385$ (the absolute magnitudes would be brighter by 4.5 mag). If we take combinations of luminosity and density evolutions, we find that the constraint on L_{50} is dominated by p_{ρ} and similar to that determined for density evolution alone. For $p_{\rho} = 1, 2$ and 3, the absolute magnitudes are brighter by 0.7, 1.6 and 3.0, respectively.

Even allowing for density evolution, the absolute magnitudes of the host galaxies are quite small. The brightest active galactic nuclei (quasars) have absolute magnitudes of at least -23 and we rule them out with a confidence greater than 90%. Absolute magnitudes of Seyfert galaxies can range from -18 to -23 and would be consistent with the observations only if there is strong density evolution. Analyses¹⁹ of the distributions of active galaxies imply that they might show luminosity evolution (p_L) but not strong density evolution (p_{ρ}). Thus, Seyfert galaxies seem unlikely as the hosts for GRBs. However, the luminosity function for dim active galaxies is poorly known²⁰ and could diverge at low absolute magnitude such that the mean brightness is fainter than -18 magnitude. Some models²¹ use the high photon production rate of active galaxies and, given our L_0 , must occur within ~ 100 Schwarzschild radii of the central black hole. Other

models²² assume a merger of a compact binary (for example, neutron star-neutron star) and it is not clear if the merger sites scale as the density of normal stars or require massive black holes or stellar cusps. $\square$

Received 22 March; accepted 9 September 1993.

1. Meegan, C. A. et al. *Nature* **335**, 143-145 (1992).
2. Meegan, C. A. et al. in *Gamma-ray Bursts*, Proc. of the Huntsville Conference (eds Paciesas, W. & Fishman, G.) 61-69 (Am. Inst. Phys. publ. No. 265, New York, 1992).
3. Mao, S. & Paczynski, B. *Astrophys. J.* **389**, L13-L16 (1992).
4. Paczynski, B. *Astrophys. J.* **308**, L43-L46 (1986).
5. Paczynski, B. in *Gamma-ray Bursts*, Proc. of the Huntsville Conference, (eds Paciesas, W. & Fishman, G.) 144-148 (Am. Inst. Phys. publ. No. 265, New York, 1992).
6. Piran, T. *Astrophys. J.* **389**, L45-L48 (1992).
7. Wickramasinghe, W. et al. *Astrophys. J.* (submitted).
8. Dermer, C. *Phys. Rev. Lett.* **68**, 1799-1802 (1992).
9. Fenimore, E. E. et al. *Nature* **357**, 140-142 (1992).
10. Mao, S. & Paczynski, B. *Astrophys. J.* **388**, L45-L48 (1992).
11. Fenimore, E. E. et al. in *Proc. of the St. Louis Compton Gamma-ray Obs. Conference* (eds Friedlander, M., Gehrels, N. & Macomb, D.) 744-748 (Am. Inst. Phys. Publ. No. 280, New York, 1993).
12. Teegarden, B. et al. in *Proc. of the St. Louis Compton Gamma-ray Obs. Conference* (eds Friedlander, M., Gehrels, N. & Macomb, D.) 860-863 (Am. Inst. Phys. Publ. No. 280, New York, 1993).
13. Fenimore, E. E. et al. in *Gamma-ray Bursts*, Proc. of the Huntsville Conference (eds Paciesas, W. & Fishman, G.) 108-112 (Am. Inst. Phys. Publ. No. 265, New York, 1992).
14. Lampton, M., Margon, B. & Bowyer, S. *Astrophys. J.* **208**, 177-190 (1976).
15. Schaefer, B. E. in *Gamma-ray Bursts: Observations, Analyses, and Theories* (eds Ho, C., Epstein, R. & Fenimore, E. E.) 107-111 (Cambridge Univ. Press, 1992).
16. Pedersen, H. et al. *Astrophys. J.* **270**, L43-L47 (1983).
17. Osterbrock, D. E. *Rep. Prog. Phys.* **54**, 579-633 (1991).
18. Schmidt, M. J., Higdon, J. & Heuter, G. *Astrophys. J.* **329**, L85-L88 (1988).
19. Boyle, B. J., Shanks, T. & Peterson, B. A. *Mon. Not. R. astr. Soc.* **235**, 935-948 (1988).
20. Hawkins, M. R. S. & Veron, P. *Mon. Not. R. astr. Soc.* **260**, 202-208 (1983).
21. Epstein, R. I., Fenimore, E. E., Leonard, P. J. T. & Link, B. in *Proc. of the Texas/PASCOS Conf.* (eds Akenlof, C. W. & Srednicki, M. A.) **685**, 565 (New York Acad. Sci.).
22. Meszaros, P. & Rees, M. J. *Astrophys. J.* **397**, 570-575 (1992).

ACKNOWLEDGEMENTS. We thank M. Briggs, A. Harding, R. Nemiroff, and B. Schaefer for conversations. This work was done under the auspices of the US Department of Energy and was funded in part by the GRO Guest Investigator Program.

Raman spectra of size-selected silicon clusters and comparison with calculated structures

E. C. Honea*, A. Ogura*†, C. A. Murray*,
Krishnan Raghavachari*, W. O. Sprenger*,
M. F. Jarrold† & W. L. Brown*

* AT&T Bell Laboratories, Murray Hill, New Jersey 07974-0636, USA

† Department of Chemistry, Northwestern University, Evanston, Illinois 60208, USA

SMALL clusters of silicon, containing between 2 and 100 atoms, have been studied extensively both because of their intrinsic interest from the point of view of chemical structure and bonding and because of the potential technological applications of cluster-assembled materials¹⁻³. *Ab initio* quantum-chemical calculations predict that very small clusters should have structures markedly different from that of the crystalline phase⁴⁻¹². Experiments on two- (ref. 13), three- (ref. 14) and four-atom¹⁵ clusters have allowed some comparison with theoretical predictions, but structural information on larger clusters has been only indirect¹⁶. Here we report on structural studies of size-selected Si_4 , Si_6 and Si_7 clusters prepared and isolated by low-energy deposition into a solid nitrogen matrix. Surface plasmon-polariton enhanced Raman spectroscopy¹⁷⁻²⁰ yields well resolved vibrational spectra for each of these clusters in which the vibrational frequencies agree well with those predicted for optimized structures calculated by *ab initio* methods. We confirm that Si_4 is a planar rhombus, and find that Si_6 is a distorted octahedron and Si_7 a pentagonal bipyramid.

† Permanent address: NEC, Microelectronics Research Laboratories, Tsukuba, Japan.

TABLE 1 Vibrational parameters for silicon clusters

Cluster	Experimental frequency (cm ⁻¹)	Calculated frequency (cm ⁻¹)*	Absolute deviation (cm ⁻¹)	Symmetry assignment	Calculated intensity†
Si ₄	345	337	8	A _g	1.0
		440		B _{3g}	0.5
	470	463	7	A _g	5.0
Si ₆ ‡	252	209	43	B _{2g}	1.0
	300	298	2	A _{1g}	2.6
	386	376	10	B _{1g}	0.6
	404	425	21	E _g	2.0
	458	457	1	A _{1g}	7.7
Si ₇	289	300	11	E ₂	1.0
	340	339	1	E ₂	0.5
	340	346	6	E ₁	0.5
	358	352	6	A ₁	1.6
	435	441	6	A ₁	4.1

* MP2/6-31G* frequencies scaled down uniformly by 5%; see text.

† Relative intensities at the HF/6-311 + G* level of theory²³ for a 90° scattering geometry, assuming no electronic resonance effects.

‡ The B_{2g} mode in Si₆, which has the largest deviation, involves the vibrations of the equatorial bonds which are unusually long (2.73 Å). Its calculated low value suggests that the extent of compression of the octahedron is probably overestimated at the MP2/6-31G* level of theory. A smaller distortion of the octahedron would increase the value of the B_{2g} mode and decrease the value of the E_g mode, bringing both into better agreement with experiment.

The silicon cluster ion source and deposition apparatus have been previously described^{21,22}. Cluster ions are produced by laser vaporization, focused into a low-energy ion beam in ultra-high vacuum and size-selected in a quadrupole mass spectrometer. The ions are slowed to 25 V and co-deposited with an N₂ matrix onto a liquid-helium-cooled substrate at ~15 K. Electrons from a hot filament neutralize the cluster ions in the matrix. Cluster ion currents ranged from ~400 pA for Si₄⁺, to ~800 pA for Si₆⁺ in a spot about 3 mm in diameter.

The substrate configuration and scattering geometry are shown in Fig. 1a. The substrate is a hemicylindrical sapphire prism, with a silver film deposited *in situ* on the flat surface. A large increase in the local electric field intensity occurs when the wavevector of the surface plasmon mode on the matrix side of the silver film is matched to that of the incident laser¹⁷⁻²⁰. Figure 1b shows the N₂ matrix Raman signal versus the incident angle of the laser beam. The N₂ Raman signal was used to optimize the excitation and light collection configuration in order to detect the approximately two orders of magnitude weaker cluster Raman signals. As the enhanced intensity decreases exponentially from the Ag interface with a calculated decay length of 70 nm, the matrix and silicon clusters are deposited on the Ag film to a thickness of 70–150 nm, followed by a capping layer of pure N₂ several microns thick.

Figure 2 shows spectra for Si₄, Si₆ and Si₇ together with predicted ground-state structures and their Raman-allowed vibrational frequencies from *ab initio* calculations. In these calculations, the vibrational properties of several isomers of each cluster were initially determined at the Hartree-Fock level with the polarized double-zeta 6-3G* basis set²³. The structures and frequencies of the ground state isomers were then determined with the inclusion of electron correlation effects at second-order Møller-Plesset perturbation theory (MP2)²³. Table 1 contains the experimental frequencies along with the calculated MP2/6-31G* harmonic frequencies scaled down uniformly by 5% to account for the expected systematic overestimation²⁴ (due to anharmonicity, for example), their symmetries, and relative intensities for a 90-degree scattering geometry.

The predicted ground state structure of Si₄ is a planar rhombus (D_{2h} symmetry) with a distorted tetrahedral isomer (D_{2d}) lying about 1 eV higher in energy. The rhombus has three allowed Raman lines at 337 (A_g), 440 (B_{3g}), and 463 (A_g) cm⁻¹. The frequencies of the two A_g modes are in excellent agreement

with the experimental values of 345 and 470 cm⁻¹. The B_{3g} vibration has not been observed experimentally, probably because it is an order of magnitude weaker (Table 1) than the nearby strong A_g mode. The distorted tetrahedral isomer, on the other hand, has five allowed Raman lines at 175, 195, 420, 499 and 648 cm⁻¹ that do not agree with the observed spectrum. Together with the previous agreement between the theoretical and experimental electron affinities and excited states^{14,15,25}, the rhombus structure of Si₄ can now be considered definitive.

The predicted ground state of Si₆ is a distorted octahedron, although the nature of the distortion had not previously been definitively determined owing to the shallow nature of the potential energy surface^{6,12}. At the higher MP2/6-31G* level of theory which includes electron correlation effects, Si₆ has a tetragonal bipyramidal structure with D_{4h} symmetry (compressed octahedron). The other candidates, such as a bicapped tetrahedron and capped trigonal bipyramid, collapsed to the same D_{4h} structure on optimization. It has five allowed Raman lines at 209 (B_{2g}), 298 (A_{1g}), 376 (B_{1g}), 425 (E_g) and 457 (A_{1g}) cm⁻¹, which agree quite well with the experimental spectrum. An alternative isomer such as a hexagonal chair (microcrystalline D_{3d} structure), with allowed Raman lines at 203, 324 and 407 cm⁻¹, compares poorly with the observed frequencies. Other possible high-energy isomers^{6,12} have lower symmetry and are predicted to have many more Raman lines. We conclude that Si₆ has a distorted octahedral geometry.

Si₇ is predicted to be a pentagonal bipyramid (D_{5h} symmetry) with a tricapped tetrahedron (C_{3v}) lying about 1 eV higher in

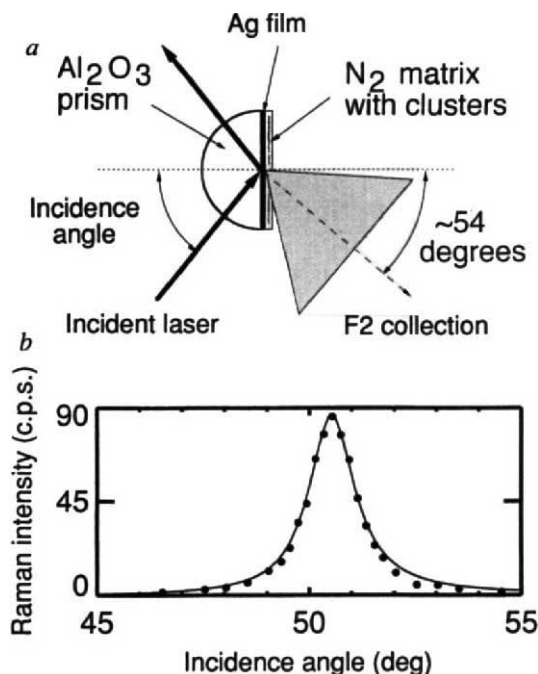


FIG. 1 a, Sample geometry and Raman optics configuration. The clusters are codeposited with N₂ on a 10-nm N₂ buffer layer on the Ag film at ~15 K in ultra-high vacuum. The excitation laser is 200 mW p-polarized with a wavelength of 488.0 nm incident at an angle ~50° to the Ag film normal, through the round side of the sapphire prism. The Raman scattered light is collected through a reentrant port with F2 collection optics at ~54° from normal on the flat side of the prism. It is focused into a triple grating spectrometer with a backthinned, liquid nitrogen cooled charge-coupled-device detector. b, Incident angle dependence of the N₂ matrix vibrational line at ~2,330 cm⁻¹, for a matrix containing, Si₄ clusters. The measured enhancement of the Raman signal at the surface plasmon-polariton resonance (~50) is consistent with Fresnel calculations using tabulated optical constants for Ag²⁶ and sapphire²⁷ and slightly modified N₂ optical constants for the cluster containing matrix, as shown by the calculated line through the data points.

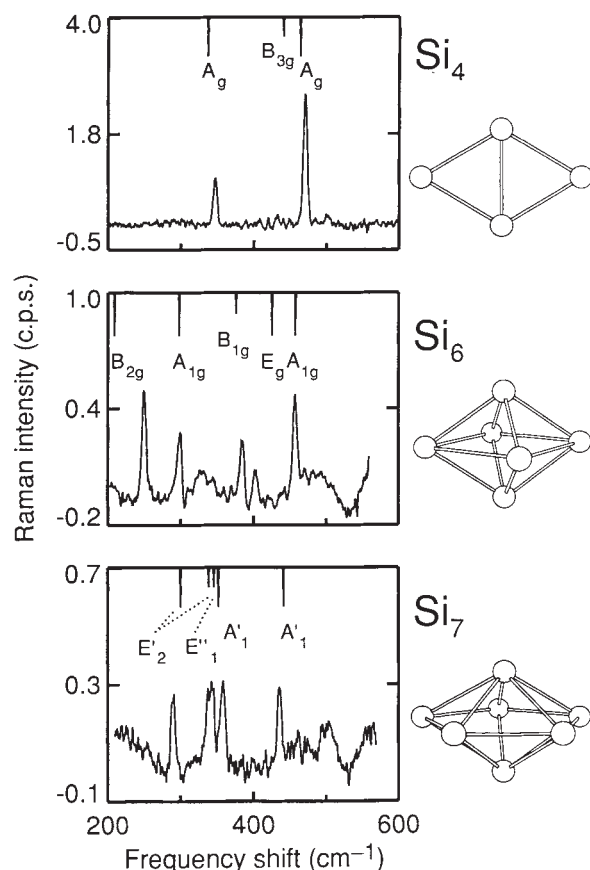


FIG. 2 Experimental Raman spectra for Si_4 , Si_6 and Si_7 deposited in a N_2 matrix, along with their predicted structures and Raman-allowed vibrational frequencies. Each experimental spectrum shown is an average of several 40-min runs and has had an average background spectrum scaled by the N_2 Raman intensity, taken 3–4 mm away from the cluster spot, subtracted from it. A slowly varying component has also been subtracted to increase the visibility of the sharp cluster lines. The predicted allowed Raman modes for the ground state structures shown at the right for each cluster are depicted as lines at the top of each experimental spectrum. Lines expected to be weaker are noted by shorter lines, although the complicated experimental geometry and surface selection rules arising from the clusters' proximity to the Ag film^{28–30}, as well as possible electronic resonances, prevent a direct comparison between the observed and calculated intensities.

energy. The D_{5h} form has five allowed Raman lines at 300 (E'_2), 339 (E'_2), 346 (E'_1), 352 (A'_1) and 441 (A'_1) cm^{-1} . Experimentally, four lines at 289, 340, 358 and 435 cm^{-1} have been observed, although the width of the line at 340 cm^{-1} suggests the possibility of two overlapping lines. The agreement between theory and experiment is excellent. The alternative C_{3v} isomer has nine allowed Raman lines (six of them $>200 \text{ cm}^{-1}$) and the agreement with experiment is much poorer. Other possible isomers with C_{3v} and C_{2v} symmetry^{6–8,12} are much higher in energy and also have many more Raman allowed lines. If these isomers exist in our samples, their concentration is roughly an order of magnitude smaller than that of the D_{5h} isomer.

Overall, 11 sharp vibrational lines have been seen for the three clusters. The mean absolute deviation between theory and experiment for the 11 lines is only about 10 cm^{-1} . This offers quite convincing evidence for the structure assignments of the calculated ground-state isomers. Although there was previous evidence from photoelectron spectra for the nature of the ground state of Si_4 (ref. 15), to our knowledge this is the first experimental evidence of the compact structures of Si_6 and Si_7 which are very different from microcrystalline forms. Further work is

in progress to elucidate the properties of larger silicon clusters and silicon materials formed by cluster deposition. □

Received 18 August; accepted 21 September 1993.

- Jarrol, M. F. *Science* **252**, 1085–1092 (1991).
- Brown, W. L., Freeman, R. R., Raghavachari, K. & Schluter, M. *Science* **235**, 860–865 (1987).
- Andres, R. P. et al. *J. Mat. Res.* **4**, 704–736 (1989).
- Raghavachari, K. *Phase Trans.* **24–26**, 61–90 (1990).
- Raghavachari, K. & Logovinsky, V. *Phys. Rev. Lett.* **55**, 2853–2856 (1985).
- Raghavachari, K. *J. chem. Phys.* **84**, 5672–5686 (1986).
- Raghavachari, K. & Rohlfing, C. M. *J. chem. Phys.* **89**, 2219–2234 (1988).
- Tomanek, D. & Schluter, M. *Phys. Rev.* **B36**, 1208–1217 (1987).
- Ballone, P., Andreoni, W., Car, R. & Parrinello, M. *Phys. Rev. Lett.* **60**, 271–274 (1988).
- Sankey, O. F., Niklewski, D. J., Drabold, D. A. & Dow, J. D. *Phys. Rev.* **B41**, 12750–12759 (1990).
- Patterson, C. H. & Messmer, R. P. *Phys. Rev.* **B42**, 7530–7535 (1990).
- Fournier, R., Sinnott, S. B. & DePristo, A. E. *J. chem. Phys.* **97**, 4149–4161 (1992).
- Arnold, C. C., Kitsopoulos, T. N. & Neumark, D. M. *J. chem. Phys.* **99**, 766–768 (1993).
- Kitsopoulos, T. N., Chick, C. J., Weaver, A. & Neumark, D. M. *J. chem. Phys.* **93**, 6108–6110 (1990).
- Arnold, C. C. & Neumark, D. M. *J. chem. Phys.* **99**, 3353–3362 (1993).
- Jarrol, M. F. & Constant, V. *Phys. Rev. Lett.* **67**, 2994–2297 (1991).
- Chen, Y. J., Chen, W. P. & Burstein, E. *Phys. Rev. Lett.* **36**, 1207–1210 (1976).
- Sakoda, K., Ohtaka, K. & Hanamura, E. *Solid St. Commun.* **41**, 393–397 (1982).
- Kurosawa, K., Pierce, R. M., Ushioda, S. & Hemminger, J. C. *Phys. Rev.* **B33**, 789–798 (1986).
- Knobloch, H., Brunner, H., Leitner, A., Aussenegg, F. & Knoll, W. *J. chem. Phys.* **98**, 10093–10096 (1993).
- Bower, J. E. & Jarrol, M. F. *J. chem. Phys.* **97**, 8312–8321 (1992).
- Honea, E. C., Kraus, J. S., Bower, J. E. & Jarrol, M. F. *Z. Phys.* **D26**, 141–144 (1993).
- Hehre, W. J., Radom, L., Schleyer, P. v. R. & Pople, J. A. *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986).
- Hout, R. F., Levi, B. A. & Hehre, W. J. *comput. Chem.* **3**, 234–250 (1982).
- Rohlfing, C. M. & Raghavachari, K. *J. chem. Phys.* **96**, 2114–2117 (1992).
- Johnson, P. B. & Christy, R. W. *Phys. Rev.* **B6**, 4370–4379 (1972).
- Handbook of Optics* (ed. Driscoll, W. G.) (McGraw-Hill, New York, 1978).
- Greenler, R. G. & Slager, T. L. *Spectrochimica Acta* **29A**, 193–201 (1973).
- Hexter, R. M. & Albrecht, M. G. *Spectrochimica Acta* **35A**, 233–251 (1979).
- Moskovits, M. *J. chem. Phys.* **77**, 4408–4416 (1982).

ACKNOWLEDGEMENTS. We thank E. Dons for RBS measurements to calibrate the silicon cluster deposition, E. Bower for developing methods and equipment construction, D. Windt for the program for Fresnel multilayer calculations, and J. Rowe, L. Brus, Y. Chabal and M. Hines for discussion.

Enhancement of axial resolution in fluorescence microscopy by standing-wave excitation

Brent Bailey, Daniel L. Farkas, D. Lansing Taylor & Frederick Lanni*

Center for Light Microscope Imaging and Biotechnology, and Departments of Biological Sciences and Physics, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213, USA

THE use of fluorescence microscopy for investigating the three-dimensional structure of cells and tissue is of growing importance in cell biology, biophysics and biomedicine. Three-dimensional data are obtained by recording a series of images of the specimen as it is stepped through the focal plane of the microscope^{1–3}. Whether by direct imaging or by confocal scanning^{4,5}, diffraction effects and noise generally limit axial resolution to about 0.5 μm . Here we describe a fluorescence microscope in which axial resolution is increased to better than 0.05 μm by using the principle of standing-wave excitation of fluorescence. Standing waves formed by interference in laser illumination create an excitation field with closely spaced nodes and antinodes, allowing optical sectioning of the specimen at very high resolution. We use this technique to obtain images of actin fibres and filaments in fixed cells, actin single filaments *in vitro* and myosin II in a living cell.

Three-dimensional imaging in a microscope requires both high transverse and axial resolution. The well-known Rayleigh for-

* To whom correspondence should be addressed.

FIG. 1 Optical configurations for standing-wave fluorescence microscopy. Fluorescence is excited in the specimen by the interference pattern of two intersecting coherent beams as described. The axis of the microscope is normal to the nodal and antinodal planes of the field. *a*, Reflected-beam optical system in which a single beam emerges on-axis ($\theta = 0$ deg) from an oil-immersed lens, passes through the specimen, and is back-reflected by a closely-apposed oil-immersed mirror. The mirror is moved axially by a piezoelectric drive, which causes an axial shift of the excitation field planes through the specimen. The optics are adjusted so that the gaussian beam leaving the objective contracts slightly to a large-diameter waist ($150\ \mu\text{m}$) at the mirror $0.1\text{--}0.2\ \text{mm}$ beyond the specimen. The standing wave field is then due to the superposition of the gaussian field $U(\mathbf{r})$ and its reflection $-U^*(\mathbf{r})$ (ref. 27). In this condition, the nodal surfaces of the unperturbed standing wave field are flat to better than 1 part in 20,000 over the field of view, and therefore can be considered as planes parallel to the mirror. The only other equipment used along with the microscope imaging system in this case is a laser and beam expander. *b*, Dual-beam optical system with a pair of oil-immersion objective lenses positioned on opposite sides of the specimen. A prism splitter amplitude-divides an expanded gaussian beam so that flat wavefronts enter the specimen independently from each side. The two beam paths in the microscope must be matched to well within the coherence length of the light source, which is a few centimetres for an ion laser. A piezoelectric drive on a mirror in one beam path serves to adjust the phase, in this case mirror movement being a full wavelength per cycle of the standing wave field. One advantage of the two-beam system is that unaberrated wavefronts enter the specimen on both sides, compared to the mirror system where phase errors accumulate on both passes when the specimen refractive index is heterogeneous. Also, the beams can be made to cross at angles other than on-axis, which has other potential advantages.

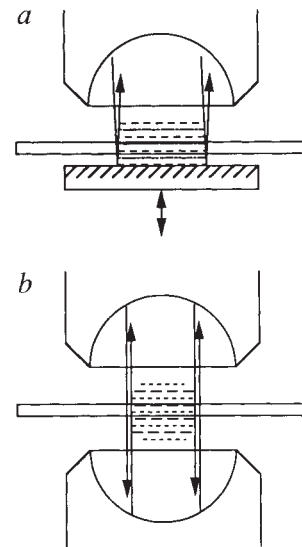
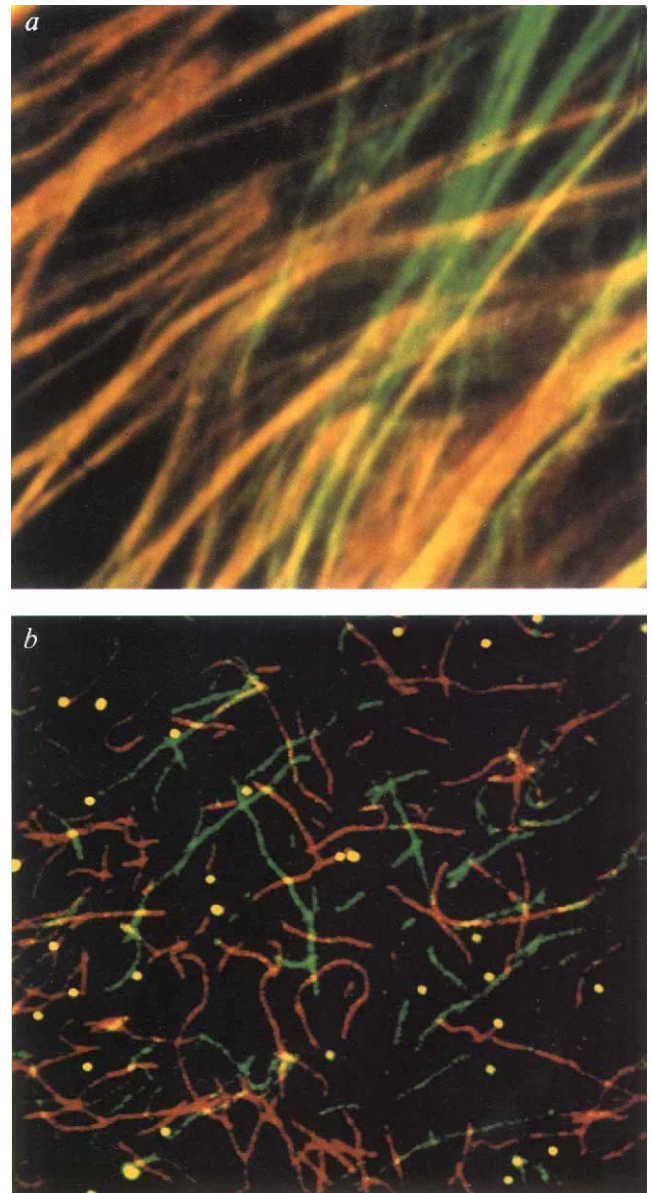


FIG. 2 Optical subsectioning of *a*, actin stress fibres in a fixed 3T3 fibroblast and *b*, stratified single actin filaments. In *a*, cells were grown on cover glasses, fixed, permeabilized, stained for F-actin with rhodamine-phalloidin, dried to cause flattening, and mounted in a thin film of a Gelvatol-water-glycerol-propyl gallate medium of high (1.45) refractive index, with a second cover glass on top. With the standing-wave field adjusted for maximum fringe flatness, different sets of overlapping stratified fibres could be made to fluoresce alternatively by shifting the fringe position. Two reflected-beam SWFM images between which the field was shifted axially $0.04\ \mu\text{m}$ were combined here as a pseudocolour overlay, and show clearly the high axial discrimination. In *b*, purified rabbit muscle actin was assembled at $0.5\ \text{mg ml}^{-1}$, labelled with rhodamine-phalloidin, and diluted to $4\ \mu\text{g ml}^{-1}$ in phosphate-buffered saline. This filament suspension was mixed with a dilute suspension of $0.23\ \mu\text{m}$ fluorescent polystyrene latex microspheres and pressed between two cover glasses to drive out all excess fluid. The microspheres set the minimum spacing between the glass surfaces, and, when viewed by conventional fluorescence microscopy, were all in focus along with F-actin trapped in the fluid layer. Movement of the standing wave field through the specimen in steps of $0.027\ \mu\text{m}$ produced large changes in relative brightness of pairs of filaments that were crossed. Two dual-beam SWFM images, between which the field was shifted by $0.081\ \mu\text{m}$, gave nearly the maximum brightness alternation, and were combined as a pseudocolour overlay. The microspheres appear as superimposed dots (yellow) in the composite image. From the ratio of background-subtracted pixel grey-scale values on the crossed filaments in either component image, we determined that the contrast ratio was 300–400%. The observed best field shift ($0.08\ \mu\text{m}$), the known node spacing ($0.19\ \mu\text{m}$), and the observation that individual filaments were of uniform brightness in each image, were consistent with all filaments being adsorbed randomly to the apposed glass surfaces at a separation of $0.27\ \mu\text{m}$. Adsorption was verified by finding regions of the specimen which were not compressed to contact, but showed all features in two distinct focal planes. Field-of-view width, $35\ \mu\text{m}$.



mula sets transverse resolution at $\sim 0.2 \mu\text{m}$ by direct imaging. In contrast, axial resolution of compact object features in fluorescence optical sectioning microscopy (FOSM) is in the range $0.7\text{--}0.9 \mu\text{m}$, which, under good circumstances, can be improved computationally to $0.4\text{--}0.5 \mu\text{m}$ ^{3,6}. Alternatively, a factor of 1.4 improvement can be achieved, both transversely and axially, by confocal scanning fluorescence microscopy (CSFM)^{7,8}, but this is often compromised owing to signal-to-noise limitations.

In standing-wave fluorescence microscopy (SWFM), two coherent plane-wave beams from a laser cross in the specimen volume, where they interfere^{9,10}. When the two fields are polarized normal to their common plane of incidence (s-polarization), are of equal amplitude, and cross at complementary angles ($\theta, \pi - \theta$) relative to the microscope axis, the resulting interference pattern has an electric field intensity that varies only axially as

$$I_{\text{ex}}(z) = 1 + \cos(Kz + \phi) \quad (1)$$

where $K = (4\pi n/\lambda) \cos \theta$, λ is the wavelength, n is the specimen refractive index, and ϕ specifies the shift of the pattern relative to the specimen. Fluorescence is excited in the specimen in proportion to $I_{\text{ex}}(z)$. The nodes and antinodes of this field, which

are planes parallel to the focal plane, have a spacing $\Delta s = \lambda/(2n \cos \theta)$. By controlling the angle θ , the node spacing can be varied down to a minimum value of $\lambda/2n$, when the two beams are counterpropagating along the axis of the microscope. By shifting the phase of one of the beams, the relative position of the field planes within the specimen can be adjusted, at constant node spacing. In our instruments, the axially periodic field is formed by the interference of two nearly-collimated gaussian laser beams propagating within the cone defined by lenses of high numerical aperture (NA) for maximum resolution and light collection efficiency (Fig. 1). Because the standing wave field diameter is comparable to the full field-of-view, images are recorded directly onto an electronic camera, rather than by scanning a focal standing wave field¹¹.

It is straightforward to estimate the axial resolution limit in SWFM. Two particles which are separated axially by half the node spacing can be illuminated alternately by shifting the node position one-half period. With blue light excitation and a specimen refractive index of 1.33 (water) or higher, $\lambda/4n$ is $0.095 \mu\text{m}$ or less. In practice, contrast alternation of specimen features was caused by field shifts equal to a quarter-fringe ($\lambda/8n$, $0.047 \mu\text{m}$) or less. This was most striking in the lamellar regions of fixed

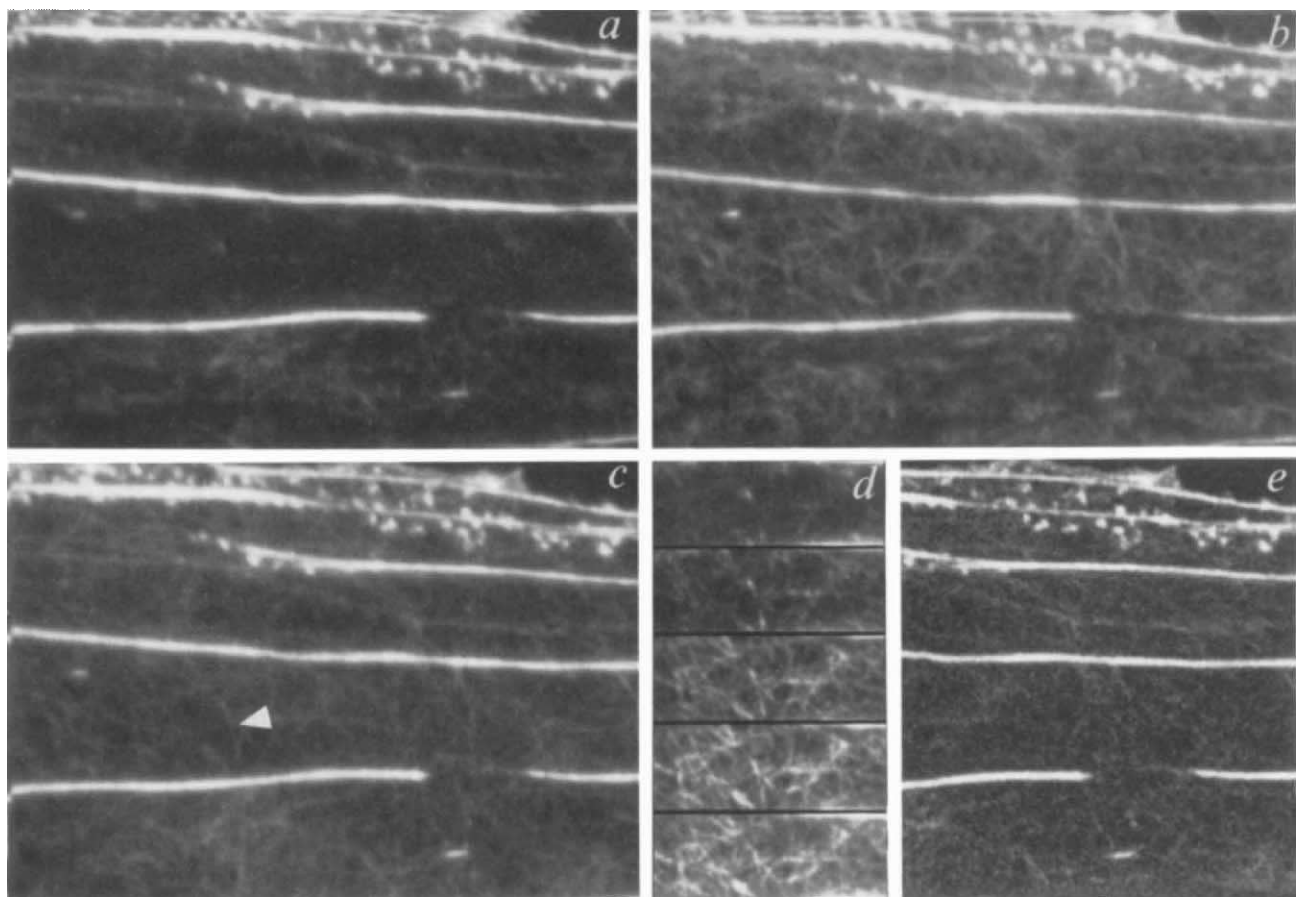


FIG. 3 Actin stress fibres and filament matrix in the lamellar cytoskeleton of a fixed quiescent 3T3 fibroblast. Cells were grown on cover glasses in complete culture medium, except that the serum concentration was reduced to 0.2% for 24 h to halt growth. The cells were then fixed, permeabilized, stained with rhodamine-phalloidin, and mounted wet to preserve 3-D structure in Gelvatol-water-glycerol-propyl gallate medium of moderate refractive index (1.40). *a* and *b*, Dual-beam SWFM images between which the field was shifted $0.036 \mu\text{m}$, showing parallel actin stress fibres and an actin filament matrix, and that the bright fibres and filaments do not occupy the same stratum. *c*, Uniform-excitation fluorescence image. With one beam blocked, the excitation field in the specimen is similar to that in conventional fluorescence microscopy, so that all structures overlap in one image. *d*, Optical subsectioning series

showing an arched actin fibre visible in *a*, *b* and *c*, (arrowhead). From top to bottom, the field was advanced towards the cover glass in steps of $0.036 \mu\text{m}$. *e*, Laser scanning confocal fluorescence image from a 3-D set of the same field of view. No significant 3-D structure was resolved with a focus increment of $0.15 \mu\text{m}$. The image acquisition period for SWFM was 2 seconds per frame, and for CSFM was 33 seconds per frame. Photobleaching of the specimen scaled similarly. Although the 5 frames in *d* show smooth structural change along the arched fibre, the standing wave intensity (equation (1)) consists of a constant term and two quadrature terms so that, rigorously, there are only three linearly independent SWFM images per focal plane. Full field-of-view width, $37 \mu\text{m}$.

index-matched fibroblasts (Fig. 2a), where stratified crossed arrays of brightly fluorescent actin fibres could be alternately imaged within one focal plane by varying only the position of the antinode. Because the axial positions resolved are tenfold smaller than the depth-of-field normally obtained in fluorescence microscopy, we call this method 'optical subsectioning'. Specifically, in thin specimens that fall entirely within the depth-of-field of a high NA objective lens and have a thickness $\leq \lambda/4n$, axial structure can be observed purely by calibrated movement of a standing wave field plane through the specimen, with no mechanical refocusing. All image features are in-focus and diffraction limited, but selectively excited. High-contrast visual discrimination of relative axial position results simply from the selective nulling of structures coincident with a nodal plane.

The fundamental definition of axial resolution used here is the limit at which two point objects separated only axially can be determined not to be located in the same transverse plane in the specimen. This necessarily involves consideration of contrast because of the absence of transverse displacement between the superposed image contributions. In a reconstituted test specimen, single fluorescent actin filaments separated axially by the diameter of 0.23- μm polystyrene latex particles could be alternately imaged with a contrast ratio exceeding 300% (Fig. 2b). In comparison, the contrast 'dip' at the classical axial resolution limit¹² of 0.7–0.9 μm is only 23%.

Constitutive actomyosin fibre assembly and transport from the peripheral cytoplasm of non-cycling Swiss 3T3 fibroblasts is a well-characterized process under study as a model of cytoskeletal self-assembly¹³. The central observation is the appearance of a new actin/myosin II fibre at the very margin of the highly spread cell, followed by movement of the fibre away from the margin towards the perinuclear region where disassembly occurs. As this process is cyclic, the lamellar cytoskeleton takes on the appearance of parallel fibres with intervening regions of $\sim 5 \mu\text{m}$ width having an alternate structure. SWFM images of cells grown under low-serum conditions, then fixed and stained (Fig. 3), show that the inter-fibre cytoskeleton is a three-dimensional matrix of randomly oriented actin filaments in a thin stratum. The optical subsectioning image series also shows that the filament matrix is located closer than the parallel fibres to the substratum, suggesting that fibre transport is a dorsal surface process in adherent cells. Neither the actin filament matrix depth

nor the axial asymmetry of filaments and fibres was apparent under uniform excitation of fluorescence or via CSFM. In addition to improved axial resolution, individual filaments were more visible in SWFM owing to nulling out of overlapping features, even though, in principle, the transverse resolution of SWFM is no different than in FOSM.

Living-cell specimens provide the greatest opportunity and challenge in application of SWFM. Active motion of cytoskeletal elements and organelles at a typical rate of $0.1 \mu\text{m s}^{-1}$ in mammalian cells¹⁴ sets the requirement that three-dimensional image data be recorded quickly if they are to be self-consistent. Likewise, the optical radiation dose received by live specimens must be minimized. As SWFM is a direct-imaging, selective-excitation technique, it has both speed and efficiency advantages over CSFM. However, little can be done to reduce the refractive index heterogeneity^{15–17} of the interior of a living cell, and the resulting light scattering or wavefront distortion could complicate the standing wave field. In living growth-phase 3T3 fibroblasts containing a fluorescent analogue of smooth-muscle myosin II, selective excitation of stratified structures was still obvious (Fig. 4), even in thicker regions requiring strictly multiple focal-plane SWFM image sets. Only sparse, small myosin aggregates were seen in the marginal cytoplasm, a zone known to be dense with newly assembled actin filaments^{14,18–21}. In the lamellum, larger myosin aggregates were numerous, and dispersed through a layer $\sim 0.2 \mu\text{m}$ thick.

Enhanced axial resolution in SWFM is due to excitation field structure, and therefore not directly limited by diffraction. In mathematical terms, standing wave excitation is equivalent to multiplying the point-spread function by $I_{\text{ex}}(z)$, in turn equivalent to shifting the optical transfer function axially in reciprocal space by a distance equal to the spatial frequency of the standing wave field. This permits recovery of information that is absent or very heavily attenuated under uniform excitation. The optical transfer function can be shifted axially by as much as $4\pi n/\lambda$, giving a fivefold improvement in axial resolution. If in addition, the specimen is very thin, an improvement of an order of magnitude may be gained through optical subsectioning. To make full use of SWFM for three-dimensional imaging of whole cells, both focal plane and nodal plane control must be integrated in the instrument. For specimens thicker than those used here, computational deconvolution^{3,22–26} should yield significant improvement in SWFM images, as it does in FOSM.

Beyond cellular specimens, we expect that there will be many applications of SWFM in the study of macromolecular or organelle assembly, particularly in cases where a transparent substratum can be used to hold structures in a known orientation for optical subsectioning. Examples include supported or unsupported lipid membranes, microtubule-based assemblies, and chromosomes. The spatial precision of experiments in which light-activated or 'caged' reagents are used may be improved in some cases by standing-wave illumination. $\square$

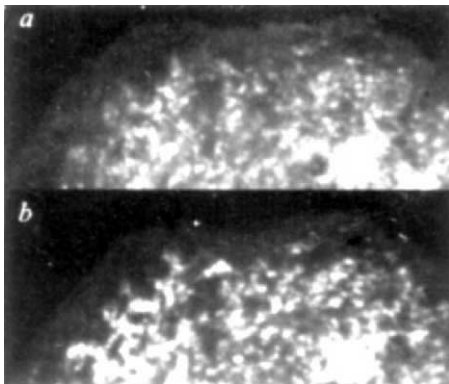


FIG. 4 Fluorescent analogue of myosin II in living 3T3 fibroblast. Cells were grown on coverglasses and microinjected with the rhodamine analogue of chicken gizzard myosin II as described in ref. 18. Specimens were assembled by pressing a coverglass over the cells in culture medium, with a ring-shaped gasket of 12.5 μm plastic film as a spacer. The thin sandwich was then sealed against evaporation and immediately transferred to the microscope. The uniform-excitation fluorescence image (a) shows a confusion of myosin-containing structures in the lamellum. The myosin aggregates are much more clearly imaged in the SWFM image (b). Field-of-view width, 33 μm .

Received 30 July; accepted 9 September 1993.

- Fay, F. S., Fujiwara, K., Rees, D. D. & Fogarty, K. E. *J. Cell Biol.* **96**, 783–795 (1983).
- Agard, D. A. *Rev. Biophys. Bioengn* **13**, 191–219 (1984).
- Agard, D. A., Hiraoka, Y., Shaw, P., & Sedat, J. W. *Meth. Cell Biol.* **30**, 353–377 (1989).
- Wilson, T. & Sheppard, C. J. R. *Theory and Practice of Scanning Optical Microscopy* (Academic, London, 1984).
- Pawley, J. (ed) *The Handbook of Biological Confocal Microscopy* (IMR, Madison, 1989).
- Carter, K. C., et al. *Science* **259**, 1330–1335 (1993).
- Sheppard, C. J. R. & Choudhury, A. *Optica* **24**, 1051–1073 (1977).
- Brakenhoff, G. J., Blom, P. & Barends, P. *J. Microsc.* **117**, 219–232 (1979).
- Lanni, F. in *Applications of Fluorescence in the Biomedical Sciences* (eds Taylor, D. L. et al.) 505–521 (Liss, New York, 1986).
- Lanni, F., Taylor, D. L. & Waggoner, A. S. US Patent No. 4,621,911 (1986).
- Hell, S. & Stelzer, E. H. K. *J. Opt. Soc. Am. A* **9**, 2159–2166 (1993).
- Born, M. & Wolf, E. *Principles of Optics* 6th edn 439–441 (Pergamon, New York, 1980).
- Giuliano, K. A. & Taylor, D. L. *Cell Motil. Cytoskel.* **16**, 14–21 (1990).
- Fisher, G. W., Conrad, P. A., DeBlasio, R. L. & Taylor, D. L. *Cell Motil. Cytoskel.* **11**, 235–247 (1988).
- Lanni, F., Waggoner, A. S. & Taylor, D. L. *J. Cell Biol.* **100**, 1091–1102 (1985).
- Bereiter-Hahn, J., Fox, C. H. & Thorell, B. *J. Cell Biol.* **82**, 767–779 (1979).
- Ross, K. F. A. & Gordon, R. E. *J. Microsc.* **128**, 7–21 (1981).

18. DeBiasio, R. L., Wang, L.-L., Fisher, G. W. & Taylor, D. L. *J. Cell Biol.* **107**, 2631–2645 (1988).
19. Small, J. V. *J. Cell Biol.* **91**, 695–705 (1981).
20. Wang, Y. -L. *J. Cell Biol.* **101**, 597–602 (1985).
21. Fan, J., Mansfield, S. G., Redmond, T., Gordon-Weeks, P. R. & Raper, J. A. *J. Cell Biol.* **121**, 867–878 (1993).
22. Holmes, T. J. *J. Opt. Soc. Am. A* **5**, 666–673 (1988).
23. Carrington, W. A. *Soc. Photo-opt. Instrumentation Engng Proc.* **1205**, 72–83 (1990).
24. Podilchuk, C. I. & Mammon, R. J. *J. Opt. Soc. Am. A* **7**, 517–521 (1990).
25. Koshy, M., Agard, D. A. & Sedat, J. W. *Soc. Photo-opt. Instrumentation Engng Proc.* **1205**, 64–71 (1990).
26. Preza, C., Miller, M. I., Thomas, L. J. Jr., & McNally, J. G. *J. Opt. Soc. Am. A* **9**, 219–228 (1992).
27. Kogelnik, H. & Li, T. *Proc. Instn Electl Electron. Engrs* **54**, 1312–1329 (1966).

ACKNOWLEDGEMENTS. This research is supported by the NSF directly, and through the Science and Technology Centers program. We thank D. A. Pane for assistance with computerized instrument control and image acquisition, and J. Montibeller, R. DeBiasio and Y. H. Wang for help with the preparations of labelled proteins and cells.

Stabilization of the membrane protein bacteriorhodopsin to 140 °C in two-dimensional films

Yi Shen^{*†}, Cyrus R. Safinya^{*†}, Keng S. Llang[†],
A. F. Ruppert[†] & Kenneth J. Rothschild[§]

^{*} Materials and Physics Departments, and the Materials Research Laboratory, University of California, Santa Barbara, California 93106, USA

[†] Exxon Research & Engineering Co., Annandale, New Jersey 08801, USA

[§] Physics and Physiology Departments, Boston University, Boston, Massachusetts 02215, USA

[‡] To whom correspondence should be addressed

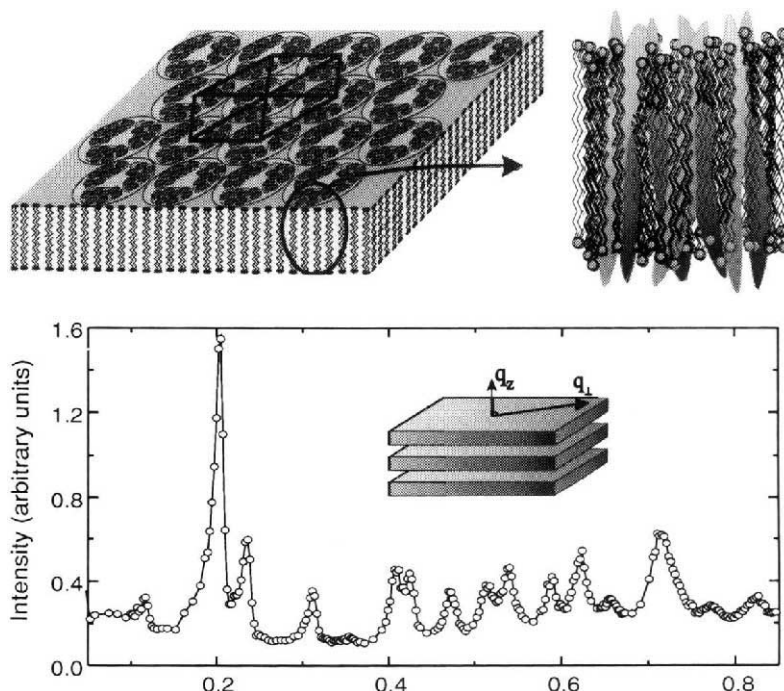
TWO-DIMENSIONAL assemblies of membrane proteins (see ref. 1, for example) such as bacteriorhodopsin are of current interest because of their potential application in technological areas as diverse as molecular electronics and optical switching², molecular sieves^{3,4} and the lithographic fabrication of nanometre-scale

FIG. 1 Top, Schematic top view of the hexagonal lattice of the purple membrane composed of bacteriorhodopsin (bR) trimers with a blow-up of a side view (redrawn from refs 14, 15) of one bR molecule showing the seven α -helices that span the membrane. The purple colour resulting from the absorption of light by the purple membrane, which in its light-adapted form has an absorption maximum near 570 nm, is due to retinal covalently attached to bR (not drawn), found also in the visual pigment rhodopsin^{7–9}. High-resolution cryoelectron diffraction²⁵ and spectroscopic techniques²⁶ have led to a more detailed 3-dimensional structure of bR and its retinal chromophore. Bottom, A typical X-ray intensity scan in reciprocal space along a direction parallel to the membrane plane through the many ordered peaks ((1, 0) through (4, 4)) of the 2-dimensional hexagonal lattice at 27 °C and relative humidity 60%. The inset shows the scattering geometry with respect to the multilayer orientation. Purple membrane was prepared from *Halobacterium halobium* R1M1 (ref. 27). A suspension of lyophilized purple membrane in distilled water (5 mg ml⁻¹) was centrifuged and 1 mg was spread on a thin 25- μ m (hydrophilically prepared) silicon wafer transparent to X-rays and covering an area of diameter about 10 mm. The wet membrane/substrate preparation was then placed in a closed container and brought into equilibrium with deionized water at 100% relative humidity. This sample procedure yielded highly oriented $\sim$ 10- μ m-thick multilayer samples with a mosaic spread of the layer normals of $\sim$ 15 deg measured through a standard crystallographic 'rocking curve' X-ray scan. The X-ray scattering experiments were carried out using both an 18 KW rotating anode X-ray generator, and at higher

patterns^{5,6}. Here we report that bacteriorhodopsin^{7–9} can retain its folded native structure to temperatures as high as 140 °C when incorporated in multilayer structures of self-assembled, ordered films. Synchrotron X-ray scattering reveals that, under hydrated conditions, the two-dimensional lattice in multilayer films exhibits a reversible solid-liquid transition at about 69 °C, followed by irreversible denaturing of the bacteriorhodopsin at about 90 °C. But in dry films the melting transition and denaturation are suppressed up to 140 °C. These results suggest that it may be feasible to use multilayer assemblies of functional proteins and enzymes^{10,11} in high-temperature applications.

The pioneering work of Blaurock, Oesterhelt and Stoekenius^{7–9} in the early 1970s led to the elucidation of the biological function of the integral membrane protein bacteriorhodopsin (bR). This protein functions as a light-driven proton pump and is involved in photosynthesis carried out by *Halobacterium halobium*, an extreme halophile (living in high salt environments) and a member of Archaeobacteria^{12,13} which thrive under conditions inhospitable to common eubacteria. Archaeobacteria also include the high-temperature extreme thermophiles¹², which undergo optimal growth at around 90 °C and are stable to about 110 °C (ref. 12). To our knowledge, the structural stability of self-assembled bR at 140 °C, which we report here, is unprecedented for a folded protein, including those of the extreme thermophiles¹².

Unlike water-soluble globular proteins, membrane proteins¹ can be incorporated into a two-dimensional planar lipid membrane. In its native form, bR self-assembles into a two-dimensional crystalline purple membrane which is part of the plasma membrane of the bacterium. The initial X-ray scattering work⁹ showed that the purple membrane, which has a diameter of $\sim$ 0.5 μ m and bilayer thickness of 49 Å, consists of trimers of bR which self-assemble in a hexagonal lattice with a spacing of 61.6 Å, with the remaining space filled by lipid (Fig. 1, top). Using electron microscopy, Henderson and Unwin^{14,15} showed that each bR molecule (shown schematically in Fig. 1, top right) consists of a polypeptide chain that folds through the membrane to form seven α -helices; this is consistent with infrared dichroism measurements (see ref. 16, for example).



resolution, the synchrotron source at the National Synchrotron Light Source. Details of the experimental set-up have been described^{28,29}.

Figure 1 shows a typical X-ray intensity scan as a function of $q_{\perp}$ in reciprocal space parallel to the membrane at room temperature, through the numerous sharp in-plane peaks of the ordered hexagonal lattice of the purple membrane, with lattice spacing $a=61.6$ Å. This highly developed two-dimensional ordering is in sharp contrast to the commonly occurring ordered L_{β} phases of membranes which comprise lipids for which normally only one or two low-order diffraction peaks are observed¹⁷.

The temperature-dependent behaviour of multilayer stacks of purple membrane at high relative humidity is shown in Fig. 2. A $(0, 0, q_z)$ scan (Fig. 2A) along the layer normal direction through the first six harmonics of the structure factor shows a very large interlayer spacing for the multilayer structure, with $d=2\pi/q_1=196$ Å. Figure 2B includes plots of three in-plane scans: plot (a) shows the sharp (10), (11) and (20) peaks of the ordered phase below the melting transition at $T=65.8$ °C, which are replaced by a broad peak of the liquid structure factor indicative of short-range positional order just above T_c in the disordered phase at $T\approx 69.2$ °C (plot b). Upon further heating to $T=99$ °C (plot c), no remnant of the broad liquid peak is seen in the disordered phase. The melting transition of the two-dimensional bR lattice, which we found to be fully reversible, occurs at $T_c\approx 69.0$ °C. The slight shift in the position of the liquid peak towards lower $q_{\perp}$ values is consistent with a slightly less dense disordered phase with short-range positional order of the trimers of bR which make up the ordered phase. In contrast, the disappearance of the liquid peak for temperatures at around 99 °C, its irreversible behaviour when the liquid peak does not reappear

upon cooling from this temperature, and the bleaching of the purple colour of the multilayer sample, are all consistent with the onset of denaturation of bR.

The behaviour under low-humidity conditions is entirely different and quite unexpected. The left panels of Fig. 3 show plots of the X-ray intensity under flowing nitrogen conditions, for scans along the in-plane direction (Fig. 3a), the q_z direction through the edge of the $q_{\perp}=(1, 1)$ scattering peak (Fig. 3b), and the multilayer stacking direction (Fig. 3c), both at room temperature and at $T=132$ °C. The behaviour for temperatures within this range is essentially identical. The $(0, 0, q_z)$ scans (Fig. 3c) show that the multilayer structure is unchanged from room temperature to very high temperatures at about 132 °C. The order-disorder transition, which occurs around $T=69$ °C at high humidities (Fig. 2B), is suppressed entirely under these dry conditions as the membrane layers are forced to stack closely to one another.

Figure 3d compares high-resolution plots of the X-ray intensity under vacuum, for scans along the in-plane direction at $T=76$ °C and 140 °C. The observation of sharp in-plane peaks of the $q_{\perp}$ scans of Fig. 3a and d is remarkable and shows that under dry conditions, the ordered hexagonal lattice remains intact up to temperatures as high as 140 °C. Furthermore, the fact that the X-ray structure factor intensities of Fig. 3a-d are essentially unchanged from room temperature to 140 °C indicates that bacteriorhodopsin remains in its tightly coiled α -helix conformation, with bR trimers self-assembled in the hexagonal lattice. We have observed a reduction in the degree of rehydration of multilayers previously exposed to these extreme temperatures of about 140 °C. This suggests that the membrane surface charge density associated with the hydrophilic amino-acid sections of bR protruding between layers is affected by these high temperatures. Although the self-assembled lattice is stable at these temperatures over several days, we found that at temperatures approaching 160 °C, bR denatures irreversibly over a period of a few hours, with the loss of the sharp diffraction peaks observed at 140 °C and of the purple colour.

Figure 3e shows, the visible absorption spectrum of three dry purple membrane samples measured at room temperature: curve 1 is for the sample at 23 °C ($\lambda_{\max}$ 560 nm), whereas the other two samples were cycled through 140 °C (curve 2, $\lambda_{\max}$ 567 nm), which resulted in a reduced intensity, broadening and slight red shift, and 160 °C (curve 3), where the sample bleached irreversibly and showed a flat visible band. The retention of the visible absorption at ~ 570 nm of all the dehydrated samples heated up to 140 °C contrasts sharply with hydrated purple membrane and

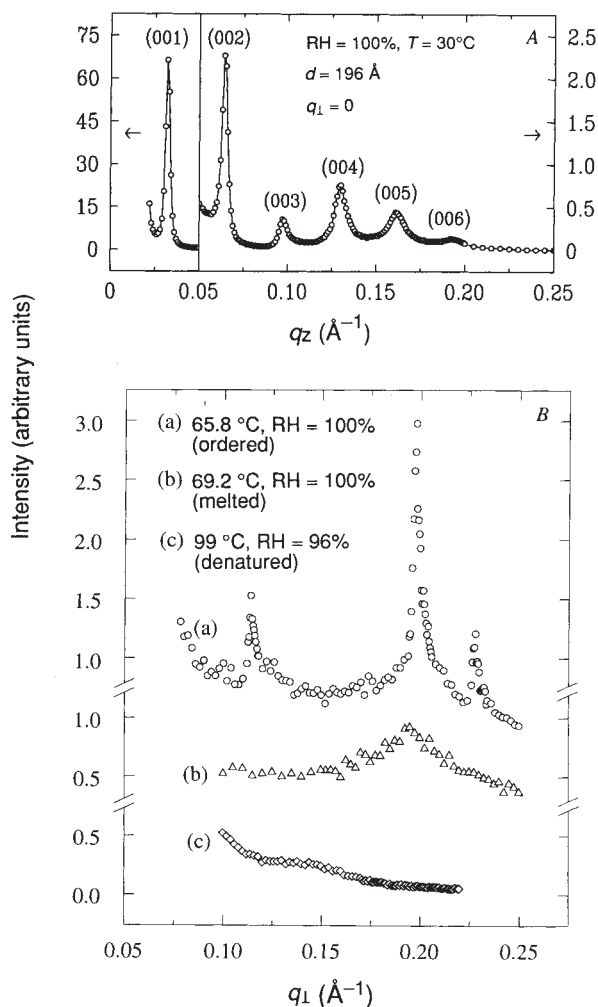
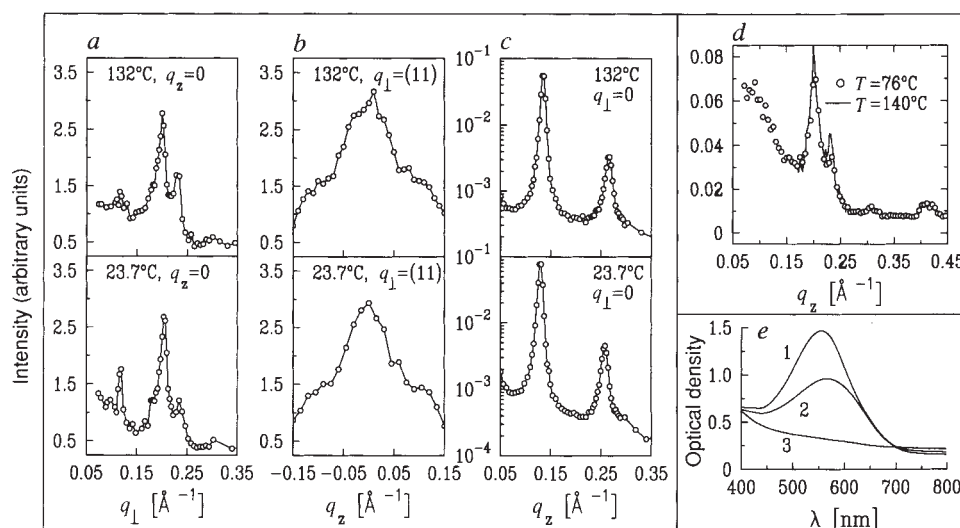


FIG. 2 High-resolution X-ray intensity scans on multilayers at high humidity (sealed in an X-ray capillary containing deionized water). A, a $(0, 0, q_z)$ scan normal to the layers through the first six harmonics of the structure factor of the multilayer, which shows the large interlayer spacing with $d=2\pi/q_1=196$ Å at 100% relative humidity (RH). B, In-plane scans parallel to the membrane plane: (a) the ordered sharp peaks below the melting temperature of the 2-dimensional hexagonal lattice ($T_c=69$ °C); (b) the broad peak of the liquid structure factor of the disordered phase just above T_c ; and (c) the absence of the liquid peak at higher temperatures when bacteriorhodopsin (bR) begins to denature and the membrane purple colour bleaches. Once denaturation (the coil-random transition of the α -helices) has set in, the behaviour is irreversible upon cooling. The broad peak around 0.13 Å⁻¹ is from the tail of the $(0, 0, 1)$ stacking peak which extends down to the $q_{\perp}$ plane because of the very large misorientation of the multilayer under these high humidity and temperature conditions. The scattering profiles are vertically shifted for clarity. Temperature control was about ± 0.1 °C. These data are consistent with differential-scanning calorimetry studies of purple membrane dispersed in water basal salts or phosphate buffers²³, which indicated a transition of ~ 70 °C associated with the lattice melting²⁴, followed by a broad denaturation transition starting at ~ 90 °C.

FIG. 3 Low-humidity (dry), low- and high-temperature data from multilayers of bacteriorhodopsin (temperature control was about $\pm 0.1^\circ\text{C}$ for nitrogen-dried samples, and $\pm 0.5^\circ\text{C}$ for samples under vacuum). *a*–*c*, X-ray intensity scans under flowing dry nitrogen at 23.7 and 132°C : *a*, in-plane scans through the first four peaks ((10), (11), (20) and (21)) of the hexagonal lattice; *b*, q_z scans through the $q_\perp = (11)$ peak, showing a width much broader than those of the in-plane scan (*a*), which is indicative of the lack of interlayer positional correlations and arises entirely from the molecular form factor. This suggests that the suppression of the in-plane melting transition is not a result of the lock-in of interlayer positional correlations of neighbouring hexagonal lattices and a cross-over from 2-dimensional to 3-dimensional behaviour; *c*, (0, 0, q_z) scans along the layer normal are indicative of the well-defined multilayer nature of the samples, with interlayer spacing $d = 48.5\text{ \AA}$ at 23°C and $46.4\text{ \AA}$ at 132°C . *d*, X-ray intensity scans along the in-plane direction under vacuum ($P \approx 10^{-8}$ torr) conditions at 76 and 140°C . Because of the sample holder configuration in our vacuum chamber, scans were recorded with the incident beam at grazing angles to the multilayer film. This resulted in the excess small-angle scattering originating from the substrate which swamped the (10) peak but allowed us to see the strong (11), (20), (21) and higher order peaks of the hexagonal lattice clearly. The sharp in-



plane peaks (*a* and *d*) result from the highly ordered hexagonal lattice of bacteriorhodopsin which retains its structure to these very high temperatures ($T = 140^\circ\text{C}$). *e*, Visible absorption spectrum of three dry purple membrane samples measured at room temperature. Curve 1: sample was held at 23°C ; curve 2: sample cycled through and maintained at 140°C for 2 d; curve 3: sample cycled through and maintained at 160°C for 4 h, which resulted in bleaching. Samples were exposed to room light and then measured in a Varian 2390 spectrophotometer; measuring time was 45 min.

with bacteriorhodopsin reconstituted in various lipids and detergents, which exhibit an irreversible structural denaturation below 100°C , that is characterized by a loss of visible absorption^{18–20}. The proton-pumping activity of bacteriorhodopsin has not been measured in the high-temperature films, but the retention of the purple colour at 140°C , together with the X-ray data, is strong evidence that the overall structure of the protein, the protonated Schiff-base bond attaching the retinal chromophore to the lysine residue at position 216, and the protein–chromophore interactions^{2,7–9}, are all maintained. The partial red shift and broadening in the absorption of the high-temperature films is not unexpected and may be due to a shift in equilibrium of bR towards the ‘O’ intermediate, the last in the bR photocycle², which absorbs near 640 nm. A similar effect has been observed in purple membrane when the temperature is raised to 35°C (ref. 21) and also in the mutant Tyr 185→Phe at room temperature²².

Although it is not clear how the removal of water, which results in the forced stacking of the membrane sheets and the hindrance of out-of-plane fluctuations, suppresses the temperature-induced order–disorder transition in purple membrane, it is likely that the absence of the melting transition enhances the stability of the protein structure at these unusually high temperatures around 140°C . Our data suggest that the dehydrated two-dimensionally ordered phase of bR results in more extensive inter-protein interactions and a larger pressure of the ordered phase up to 140°C , relative to that of the corresponding liquid phase of bR found by us and others^{23,24} above 70 – 80°C in hydrated multilayers. This then enhances the bacteriorhodopsin folding forces by exerting steric constraints that prevent protein structural changes, which otherwise would disrupt the bR α -helical bundle at these high temperatures. This is in agreement with previous results^{18,20} which indicate that elimination of the crystalline lattice lowers the thermal stability of bR.

Aside from the direct potential applications to bacteriorhodopsin, we believe that our finding should lead to the stabilization of the structure of many other proteins and enzymes: that is, by artificially preparing multilayers of ordered arrays of

macromolecules^{10,11}, significant high-temperature stability should be generated. Such stability will allow the investigation of the structure of a folded protein under these extreme temperature conditions. □

Received 12 July; accepted 21 September 1993.

1. *Electron Microscopy at Molecular Dimensions: State of the Art and Strategies for the Future* (eds Baumeister, W. & Vogell, W.) (Springer, Berlin and New York, 1980).
2. Birge, R. R. *Rev. Phys. Chem.* **41**, 683–733 (1990).
3. Sleytr, U. B. & Sara, M. *Appl. Microbiol. Biotechnol.* **25**, 83–90 (1986).
4. Baumeister, W. & Lembecke, G. *J. Bioenerg. Biomembr.* **24**, 567–575 (1992).
5. Douglas, K., Clark, N. A. & Rothschild, K. J. *Appl. Phys. Lett.* **58**, 692–694 (1990).
6. Douglas, K., Devaud, G. & Clark, N. A. *Science* **257**, 642–644 (1992).
7. Oesterheld, D. & Stoekenius, W. *Nature New Biol.* **233**, 149–152 (1971).
8. Oesterheld, D. & Stoekenius, W. *Proc. Natn. Acad. Sci. U.S.A.* **70**, 2853–2857 (1973).
9. Blaurock, A. E. & Stoekenius, D. *Nature New Biol.* **233**, 152–155 (1971).
10. Uzgiris, E. E. & Kornberg, R. D. *Nature* **301**, 125–129 (1983).
11. Blankenburg, R., Meller, P., Ringsdorf, H. & Saless, C. *Biochemistry* **28**, 8214–8221 (1989).
12. Brock, T. D. & Madigan, T. in *Biology of Microorganisms* 6th edn (Prentice Hall, New Jersey, 1991).
13. Woese, C. R. *Sci. Am.* **244**, 98–122 (1981).
14. Henderson, R. & Unwin, P. N. T. *Nature* **257**, 28–32 (1975).
15. Unwin, P. N. T. & Henderson, R. *J. molec. Biol.* **94**, 425–440 (1975).
16. Braiman, M. S. & Rothschild, K. J. *A. Rev. Biophys. biophys. Chem.* **17**, 541–570 (1988).
17. Smith, G. S., Sirota, E. B., Safinya, C. R., Plano, R. J. & Clark, N. A. *J. chem. Phys.* **92**, 4519–4529 (1990).
18. Brouillette, C. G., McMichens, R. B., Stern, L. J. & Khorana, H. G. *Proteins: Struct. Funct. Genet.* **5**, 38–46 (1989).
19. Kresheck, G. C. et al. *J. Photochem. Photobiol. B* **7**, 289–302 (1990).
20. Kahn, T. W., Sturtevant, J., Engelman, D. M. *Biochemistry* **31**, 8829–8839 (1992).
21. Fukuda, K. & Koyama, T. *Biochemistry* **31**, 11740–11747 (1992).
22. Sonar, S., Krebs, M. P., Khorana, H. G. & Rothschild, K. *Biochemistry* **32**, 2263–2271 (1993).
23. Jackson, M. B. & Sturtevant, J. M. *Biochemistry* **17**, 4470–4474 (1978).
24. Hiraki, K., Hamanaka, T., Mitsui, T. & Kito, Y. *Biochim. biophys. Acta* **647**, 18–28 (1981).
25. Henderson, R. et al. *J. molec. Biol.* **213**, 899–929 (1990).
26. Rothschild, K. J., Braiman, M. S., Mogi, T., Stern, L. J. & Khorana, H. G. *FEBS Lett.* **250**, 448–452 (1989).
27. Oesterheld, D. & Stoekenius, W. *Meth. Enzym.* **31**, 667–678 (1974).
28. Safinya, C. R. et al. *Phys. Rev. Lett.* **57**, 2718–2721 (1986).
29. Safinya, C. R., Sirota, E. B., Roux, D. & Smith, G. S. *Phys. Rev. Lett.* **62**, 1134–1137 (1989).

ACKNOWLEDGEMENTS. We thank S. Sonar and N. Patel for help with purple membrane preparation; R. Bruinsma, N. Clark, P. Pincus, T. Lubensky, D. Morse, D. Nelson, and S. Sonar for discussion; and D. Thirumalai for critically reading the manuscript. The National Synchrotron Light Source (NSLS) is supported by the US Department of Energy. The Materials Research Laboratory at Santa Barbara is supported by the NSF. C.R.S. acknowledges the NSF for a biomolecular materials grant, and the Exxon Research and Engineering Company for a research grant; K.J.R. thanks the NSF and the ARO for support.

Effect of a lowered water table on nitrous oxide fluxes from northern peatlands

Pertti J. Martikainen*, Hannu Nykänen*,
Patrick Crill† & Jouko Silvola‡

* Department of Environmental Microbiology, National Public Health Institute, PO Box 95, FIN-70701 Kuopio, Finland

† Complex Systems Research Center, EOS, University of New Hampshire, Durham, New Hampshire 03824-3525, USA

‡ Department of Biology, University of Joensuu, PO Box 111, FIN-80101 Joensuu, Finland

NORTHERN peatlands contain 20–30% of the total organic nitrogen and carbon in the world's soils^{1,2}, and thus they apparently have the potential to exert a significant influence on the global atmospheric budget of the greenhouse gases carbon dioxide, methane and nitrous oxide (N₂O). In the drier, warmer summer conditions predicted at high latitudes by some climate models^{3,4} as a result of greenhouse-gas forcing, northern peatlands would become drier, increasing the rate of mineralization of organic matter^{1,5} and of the microbial processes that produce N₂O. These regions might therefore be expected to exert a strong feedback on climate. But whereas methane emissions have been well studied^{6,7}, little is known about the effect on N₂O fluxes of changes in the level of peatland water tables. Here we present a comparison of present-day N₂O fluxes from virgin peatlands in Finland with those from sites in the same regions that were drained by ditching 30 and 50

years ago. The lowered water table had no effect on N₂O emissions from nutrient-poor peat but enhanced those from nutrient-rich peat. We estimate that equivalent drying caused by climate change would increase the total emissions of N₂O from northern peatlands by 0.03–0.1 teragrams of nitrogen per year, which is just 0.3–1% of the present global annual emissions. Thus northern peatlands are unlikely to exert a significant climate feedback from N₂O emissions.

The sites examined in eastern (site 1, 62° 51' N and 30° 53' E) and central Finland (sites 2, 3, 4 and 5 which are a part of the Lakkasuo mire complex, 61° 47' N and 24° 18' E) differed in their nutrient status and vegetation (Table 1), and represent both ombrotrophic and minerotrophic peatlands. In contrast to minerotrophic peatlands, ombrotrophic peatlands receive no water from surrounding mineral soils, so atmospheric deposition is their only source of mineral nutrients. The ombrotrophic bog and minerotrophic fen sites in the Lakkasuo mire complex had both virgin areas and areas drained for forestry by ditching in 1961. Site 1 was first drained around 1940 and redrained in 1979. At the minerotrophic sites the ditching had promoted successional changes toward very different ecosystems from the original. Typical fen species were replaced by spruce swamp and forest species (Table 1). At ombrotrophic sites, botanical changes due to draining were smaller with far less tree production. Although the original depth and distance between ditches were similar at all drained Lakkasuo sites, the higher tree growth in the minerotrophic sites⁸ (Table 1) had increased evapotranspiration and contributed to the lower water table (Fig. 1a–c).

Nitrous oxide emissions at all wet, virgin sites were low. Virgin peat sites even showed some uptake of N₂O as has been observed in laboratory experiments with peat soil⁹. In contrast to the ombrotrophic sites, draining increased N₂O emission at minero-

FIG. 1 Fluctuations in N₂O emissions, water table and soil temperatures in 1991 and 1992 at the five peatland sites studied. Gas fluxes were measured by a static chamber method²⁹ every second week from spring thaw to winter freeze in 1991. In 1992 fluxes were measured 1–4 times a month. During winter 1992, site 1 was measured twice, sites 3 and 4 once. (There were 2–3 chambers at virgin and drained areas of each site). Mean fluxes and the associated standard errors are shown. Gases were analysed with a gas chromatograph equipped with an electron capture detector. The sites were equipped with boardwalks to minimize disturbance caused by sample collection.

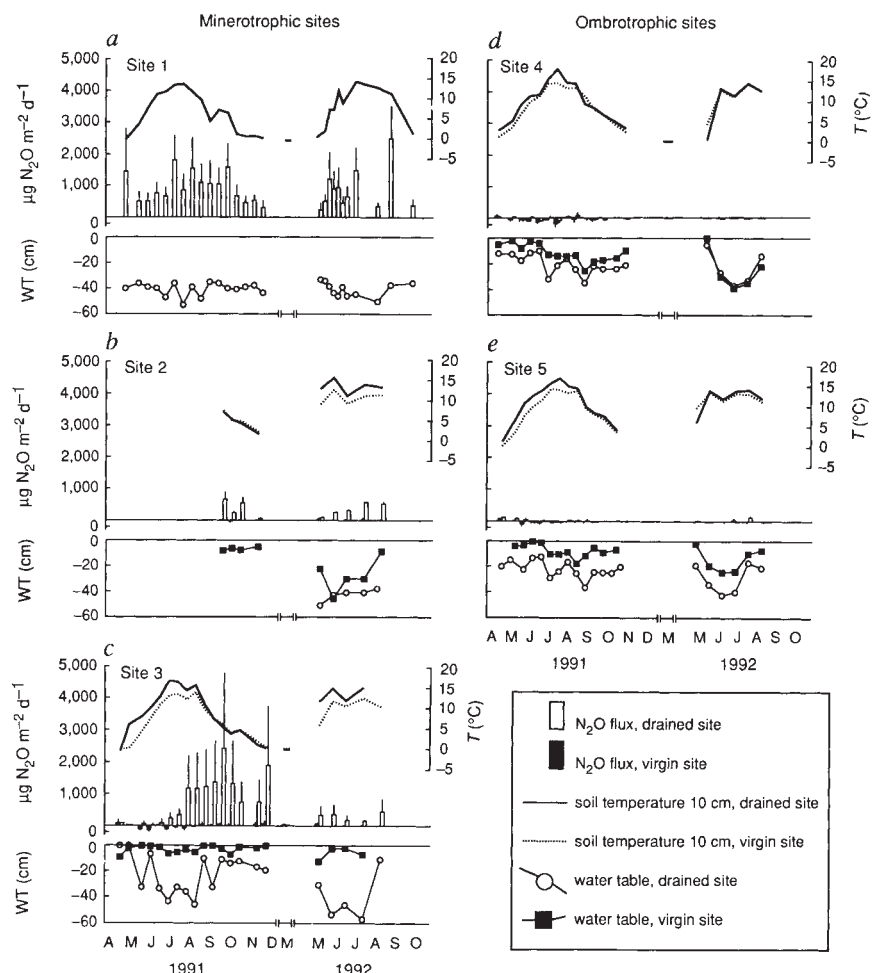


TABLE 1 The peatland sites, their botanical and nutritional characteristics, pH and annual N₂O emissions

Site number and type*	Dominating plant species in surface vegetation and changes in vegetation after drainage	Tree stand	Total N in peat (%)	pH†	N ₂ O emission‡ (g N m ⁻² yr ⁻¹)	
					1991	1992
1. Mesotrophic sedge pine mire	Drained: <i>Pleurozium schreberi</i> , <i>Polytrichum commune</i> , <i>Sphagnum angustifolium</i> , <i>Vaccinium myrtillus</i> , <i>Vaccinium vitis-idaea</i> .	160 m ³ per hectare 93% pine§ 7% birch	1.8	4.3	0.143	0.125
2. Mesotrophic flark fen	Virgin: <i>Sphagnum papillosum</i> , <i>S. angustifolium</i> , <i>S. magellanicum</i> (hummocks), <i>Trichophorum caespitosum</i> , <i>Betula nana</i> (hummocks). Drained: vegetation scarce because of abundant birch leaf litter. Much less sphagnum species than in virgin area. Typical new species: <i>Polytrichum commune</i> and <i>P. strictum</i> .	<5 m ³ per hectare 50 m ³ per hectare 95% birch 5% pine and spruce¶	n.d.	5.0	<0.004	<0.004
3. Tall sedge fen	Virgin: <i>Carex rostrata</i> , <i>Sphagnum fallax</i> , <i>S. papillosum</i> . Drained#: <i>Polytrichum commune</i> , <i>P. strictum</i> . Much less sphagnum species than in virgin area.	<1 m ³ per hectare 100% pine 90 m ³ per hectare 99% pine 1% birch	1.9	5.0	<0.004	<0.004
4. Cottongrass pine bog with <i>Sphagnum fuscum</i> hummocks	Virgin: <i>Sphagnum angustifolium</i> , <i>S. fuscum</i> , <i>Eriophorum vaginatum</i> , <i>Empetrum nigrum</i> . Drained: less sphagnum species than in virgin area. The proportion of forest mosses has increased (<i>Pleurozium schreberi</i> , <i>Polytrichum strictum</i>). The proportion of <i>Eriophorum vaginatum</i> and <i>Empetrum nigrum</i> slightly decreased.	5 m ³ per hectare 100% pine 20 m ³ per hectare 100% pine	0.5	4.3	<0.004	<0.004
5. Low-sedge bog	Virgin: <i>Sphagnum angustifolium</i> , <i>S. fuscum</i> , <i>Empetrum nigrum</i> <i>Eriophorum vaginatum</i> . Drained: clearly less sphagnum species. The proportion of hummock species has increased (<i>Pleurozium schreberi</i> , <i>Betula nana</i>)	1 m ³ per hectare 100% pine 15 m ³ per hectare 100% pine	0.4	4.2	<0.004	<0.004

n.d., Not determined.

* Sites 1, 2 and 3 are minerotrophic, 4 and 5 are ombrotrophic. Site type names according to Finnish peatland classification²⁴.

† Soil pH, measured in soil-water suspension.

‡ Annual emissions are calculated from average monthly fluxes, winter emissions are estimated from winter measurements (see measuring frequency in Fig. 1).

§ *Pinus sylvestris*.|| *Betula pendula* and *B. pubescens*.¶ *Picea abies*.

Fertilized in 1966 with 10 kg N per hectare, 6 kg P per hectare and 6 kg K per hectare. The site received 53 kg P per hectare and 63 kg K per hectare in 1984.

trophic sites (Fig. 1). Laboratory experiments with samples from peat profiles of a drained minerotrophic site showed nitrification activity, and denitrification associated with the produced nitrate (NO₃⁻) (results not shown). Low nitrification activity and lack of nitrate could explain the low N₂O production in drained ombrotrophic and virgin minerotrophic peat. It has been observed previously that nitrification activity in mires is low¹⁰⁻¹³. Although the total amount of nitrogen even in the peat profile of ombrotrophic mires is high, the concentration of nitrogen in the peat is low (<1% of the dry matter; Table 1). This low nitrogen concentration in ombrotrophic peat, together with a shortage of other nutrients such as phosphorus¹⁴, and low pH probably limit nitrogen mineralization and nitrate production which are necessary for N₂O production. The conditions for N₂O production in the forested ombrotrophic peatlands seem to be similar to those in the acid unfertilized soils of coniferous forests, which represent another important ecosystem at nor-

thern latitudes. These latter soils also have very low N₂O emission rates, 0.001–0.01 g N₂O–N m⁻² yr⁻¹ (refs 15, 16) (our emission rates for the drained minerotrophic fens were 10–100 times higher than this, see Table 1).

The N₂O emission from nitrogen-rich mire ecosystems might be expected to be much higher than that found in the present study. Fluxes of N₂O from nitrogen-rich organic soils of some disturbed meadows are higher¹⁷ than we found for the drained minerotrophic fens. Nitrogen mineralization, nitrate production and denitrification activities in disturbed meadow soils may be higher than those in soils of forested peatlands, thereby explaining the difference in the N₂O fluxes between these two ecosystems. We have also measured N₂O emissions from a drained minerotrophic fen used for agriculture that were 5–20 times higher than N₂O emissions from forested minerotrophic fens (results not shown). Cultivation practices such as ploughing and fertilization enhance availability of ammonium and nitrate for

nitrification and denitrification, thus providing good conditions for N_2O production.

There are large differences in the emission rates of N_2O in different years and seasons (site 3, Fig. 1c). The comparison of the virgin and drained fen sites shows that some lowering of the water table is required for an increase in the emissions of N_2O . However, the fluxes at drained minerotrophic fens were rather poorly correlated with the actual level of water table: the correlation with soil temperature was also weak. The N_2O emissions in late autumn and early spring could be as high as those in summer with high soil temperature (Fig. 1). The results suggest that for N_2O emissions from the northern peatlands, the development of water table would be more important than the increase in peat temperature. This same conclusion has been reached regarding the decomposition of organic matter in northern peatlands⁵. A long-term increase in temperature may affect the function of forested mire ecosystems by changing the growth rate of vegetation, and even changing the composition of tree species¹⁸. The changes in litter production associated with changes in tree stand may affect mineralization as well as the rates of nitrification and denitrification that produce N_2O in the peat profile.

There is general agreement that global warming will be greatest at northern latitudes⁴, but how this will change soil moisture is less clear¹⁹. If temperature increases without decreasing soil moisture, there will probably be no increase in N_2O emissions from northern peatlands. Whatever the future soil moisture, conditions may differ from those found in these drained peatlands, where ditching exerts a sudden and marked influence on the water table. The average water table would decrease only slowly with changing climate, and during most of year the water table would remain at a high level. This may retard the development not only of the microbiological processes but also of the tree growth which is typical for the drained minerotrophic mires. As mentioned above, trees increase evapotranspiration and thus contribute to the lowering of water table.

Significant changes in fluxes of CH_4 (refs 20–22) and CO_2 (ref. 23) have been associated with changes in water table in subarctic and boreal peatlands (the former are expected to decrease in response to drying). Our results show that the consequences for N_2O emissions are likely to be less marked. (Note, however, that some northern mires are more nutrient-rich than our minerotrophic study sites²⁴, and these fens might experience greater changes in N_2O emission rates.) About one-third of the peatlands in Finland are classified as wet minerotrophic²⁵. Although the proportion of wet minerotrophic mires is perhaps lower in the extensive peatland areas of North America and Russia than it is in Finland^{26–28}, there are at least $0.3–0.5 \times 10^6$ km² of minerotrophic mires (that correspond in type to our minerotrophic sites) in the 3.6×10^6 km² of boreal and subarctic peatlands in the world¹. If these were to dry out like our drained study sites, we estimate their average N_2O emission as $0.05–0.15$ g N_2O-N m⁻² yr⁻¹ (Table 1), giving a total of $0.03–0.1$ Tg N_2O-N yr⁻¹. This is just 0.3–1% of the present global annual N_2O emission¹⁵. □

15. Davidson, E. A. in *Microbial Production and Consumption of Greenhouse Gases: Methane, Nitrogen Oxides, and Halomethanes* (eds Rogers, J. E. & Whitman, W. B.) 219–235 (Am. Soc. for Microbiol., Washington DC, 1991).
16. Matson, P. A., Gower, S. T., Volkman, C., Billow, C. & Grier, C. C. *Biogeochemistry* **18**, 101–117 (1992).
17. Goodroad, L. L., Keeney, D. R. *J. envir. Qual.* **13** 448–452 (1984).
18. Solomon, A. M. *Oecologia* **68**, 567–579 (1986).
19. Intergovernmental Panel on Climate Change *Climate Change—the IPCC Scientific Assessment*. (Cambridge Univ. Press, 1990).
20. Roulet, N., Moore, T., Bubier, J. & Lafleur, P. *Tellus* **44B**, 100–105 (1992).
21. Moore, T. R. & Knowles, R. *Can. J. Soil Sci.* **67**, 77–81 (1987).
22. Martikainen, P. J., Nykänen, H., Crill, P. & Silvola, J. *Suo* **43**, (4–5), 237–240 (1992).
23. Silvola, J., Välijoki, J. & Aaltonen, H. *Acta for. fenn.* **191**, 1–32 (1985).
24. Ruuhijärvi, R. in *Ecosystems of the World 4B Mires: Swamp, Bog, Fen and Moor. Regional Studies* (ed. Gore, A. J. P.) 47–67 (Elsevier, Amsterdam, 1983).
25. Keltikangas, M., Laine, J., Puttonen, P. & Seppälä, K. *Acta for. fenn.* **193**, 1–94 (1986).
26. Botch, M. S. & Masing, V. V. in *Ecosystems of the World 4B Mires: Swamp, Bog, Fen and Moor. Regional Studies* (ed. Gore, A. J. P.) 95–152 (Elsevier, Amsterdam, 1983).
27. Hofstetter, R. H. in *Ecosystems of the World 4B Mires: Swamp, Bog, Fen and Moor. Regional Studies* (ed. Gore, A. J. P.) 201–244 (Elsevier, Amsterdam, 1983).
28. Zoltai, S. C. & Pollet, F. C. in *Ecosystems of the World 4B Mires: Swamp, Bog, Fen and Moor. Regional Studies* (ed. Gore, A. J. P.) 245–268 (Elsevier, Amsterdam, 1983).
29. Crill, P. M. et al. *Glob Biogeochem. Cycles* **2**, 299–327 (1988).

ACKNOWLEDGEMENTS. This research was supported by the Finnish Research Programme on Climate Change (SILMU). P. C. was also supported by the NASA Interdisciplinary Program. We thank T. Aitolaitti, U. Ahlholm, L. Halko, E. Ihanus, E. Konttinen, J. Meronen, M. Pilttonen and M. Ronkainen for field and laboratory assistance, J. Alm, K. Tolonen, H. Vasander and J. Laine for chemical and botanical data, and J. Aho for technical facilities at the Mekrijärvi Research Station.

Absence of preferred longitude sectors for poles from volcanic records of geomagnetic reversals

Michel Prévot & Pierre Camps

Centre Géologique et Géophysique, CNRS and
Université de Montpellier 2, 34095 Montpellier Cedex 05, France

ACCORDING to a recent compilation of sedimentary records from the past 12 Myr, the reversing geomagnetic field displays a marked long-term longitudinal organization which seems to be correlated with the thermal structure of the lower mantle¹. However, the crucial palaeomagnetic observation that intermediate virtual geomagnetic poles are preferentially confined to longitudes over the Americas or antipodal to this sector^{2,3} might be an artefact caused by distortion and smoothing of the geomagnetic signal by sediments^{4–6}. Here we present a compilation of ~400 intermediate poles from 121 volcanic records of excursions and reversals less than 16 Myr old, sampled at various longitudes and latitudes, which does not confirm previous inferences¹ and reinforces doubts^{4–6} about the ability of sediments to properly record a rapidly varying field. In contrast to virtual poles from excursions, those from reversals are uniformly distributed in longitude, which indicates that the reversing field is statistically axisymmetrical. Thus, we conclude that there is no evidence for control of transitional fields by the temperature distribution in the lowermost mantle.

The claim^{2,3} that virtual geomagnetic poles (VGPs) during complete or aborted reversals are preferentially confined to two sectors of opposed longitude ($120 \pm 28^\circ$ and $300 \pm 28^\circ$) is based on compilations restricted to sedimentary records. For this specific database, their claim appears statistically valid^{3,7}. What is in question is the ability of sediments reliably to record and preserve⁶ the rapid and complicated directional and intensity changes which occur when the field reverses⁸. An independent testing of the confinement hypothesis is therefore required, and lava flows provide the best opportunity. It is well established, as shown again recently⁹, that the primary remanence acquired by these rocks during their cooling does not deviate from the field direction by more than a few degrees, depending on the magnetization intensity¹⁰.

Received 15 March; accepted 9 September 1993.

1. Gorham, E. *Ecol. Appl.* **1**(2), 182–195 (1991).
2. Sjörs, H. *Int. J. Ecol. Envir. Sci.* **7**, 11–14 (1981).
3. Manabe, S. & Wetherald, R. T. *Science* **232**, 626–628 (1986).
4. Mitchell, J. F. B. *Rev. Geophys.* **27**, 115–139 (1989).
5. Oechel, W. C. et al. *Nature* **361**, 520–523 (1993).
6. Harris, R. C., Gorham, E., Sebach, D. L., Bartlett, K. B. & Flebbe, P. A. *Nature* **315**, 652–654 (1985).
7. Tyler, S. C. in *Microbial Production and Consumption of Greenhouse Gases: Methane, Nitrogen Oxides, and Halomethanes* (eds Rogers, J. E. & Whitman, W. B.) 7–38 (Am. Soc. for Microbiol., Washington DC, 1991).
8. Laine, J. & Vanha-Majamaa, I. *Ann. Bot. Fenn.* **29**, 213–233 (1992).
9. Hemond, H. F. *Ecology* **64**, 99–109 (1983).
10. Rosswall, T. & Granhall, U. *Ecol. Bull. (Stockholm)* **30**, 209–234 (1980).
11. Williams, B. L. & Wheatley, R. E. *Biol. Fertil. Soils* **6**, 141–147 (1988).
12. Rangeley, A. & Knowles, R. *Soil Biol. Biochem.* **20**, 385–391 (1988).
13. Koerselman, W., De Caluwe, H. & Kierskamp, W. M. *Biogeochemistry* **8**, 153–165 (1989).
14. Amador, J. A. & Jones, R. D. *Soil Biol. Biochem.* **25**, 793–801 (1993).

The first compilations of volcanic data, restricted to Iceland, suggested the absence of strongly preferred longitudinal sectors for transitional poles^{11,12}, but more recent analyses provide contradictory indications. Constable¹³ found that VGPs calculated from non-transitional directions < 5 Myr old are (slightly) concentrated in the same two opposite longitude sectors pointed at by sedimentary records of reversals, which she took as an indication that the transitional confinement is real. In the same vein, Hoffman¹⁴ argued (from the consideration of a set of selected transitional directions observed in some volcanic records) that two specific inclined dipolar field configurations, whose axes lie within the two preferred longitude sectors, have dominated the reversal process for the past ten million years. In contrast, Valet *et al.*¹⁵ concluded from a compilation of ten records of reversals or excursions extending in age from Miocene to Quaternary that no coherent pattern emerges from the longitudinal distribution of VGPs.

To try to solve these contradictions, we have constructed a more representative database (available from *Nature* on request, see Supplementary Information). We compiled from the palaeomagnetic literature all intermediate VGPs (defined here as having a latitude < 60°) obtained from magnetically cleaned lava flows: 121 records of excursions or reversals were thus selected. Note that in our terminology reversals are geomagnetic events resulting in a 180° change in direction, accompanied by the restoration of the initial field intensity. When these two characteristics are not observed, as for the "aborted reversals" of some authors^{14,15}, the term excursion is used. Careful reading of the selected papers enabled us to discard duplicate records of the same field direction(s) when the same geomagnetic event had been identified at different places in the same region. Similarly, because several eruptions can occur over periods of weeks or months, we considered (as previously⁸) that consecutive lava flows whose α_{95} confidence angles overlap and whose average directions do not exhibit a regular trend with time represent a single field reading. All detailed volcanic records of reversals show the absence of confinement in longitude of their generally complex VGP path^{11,16-18}. Because of this complexity, such a path cannot be described by a single datum, such as the equatorial crossing point³ (EC) or the mean value of longitude¹³ (MVL) used for sedimentary records. Commonly, several very different ECs are observed (for example three, differing by as much as 120°, for the Steens Mountain reversal¹⁶) and the VGP distribution is so far from being unimodal that calculating an average pole is meaningless. Indeed, averaging is not necessary because all VGPs represent similar time intervals (the flow cooling time),

which are here reasonably distributed over the time interval investigated. Altogether they constitute a random sampling of the transitional field by volcanic eruptions. As none of the 121 records selected here represents > 8% of the total number of VGPs, the overall distribution cannot be significantly biased by any single record. Thus, as in previous compilations of volcanic VGPs^{11,13,15}, all VGPs will be given the same weight unless otherwise specified (Fig. 2b).

According to the usual palaeomagnetic convention, VGPs correspond to south magnetic poles. Because the dynamo equations are invariant under reversal of sign of the magnetic field¹⁹, a single longitudinal singularity at the core-mantle boundary gives rise to two antipodal singularities in the longitudinal distribution of the 'palaeomagnetic' VGPs^{19,20}. Thus, we will instead use here the geomagnetic convention which consists, for a given geomagnetic event, of plotting the VGPs whose polarity is that of the pole initially located in the Northern Hemisphere. Note that when this mode of representation is used, all reversals progress from the northern towards the southern pole. Given the large size of our VGP population, the usual circular statistic tests (Hodges-Ajne, Kuiper, Watson's and Rayleigh), as described by Mardia²¹, were used to check the null hypothesis (uniformity) against either an axial (bimodal, with two antipodal modes) or a unimodal Von Mises distribution. The level of significance (risk of rejecting the null hypothesis when true) was set at 0.05. The compilation presented here covers the past 16 Myr, which is comparable with the 12 Myr coverage of sedimentary records considered in refs 3 and 15. Such durations are short compared with the time constants of lower-mantle processes²² and are therefore appropriate for unveiling a possible mantle effect. A total of 631 intermediate VGPs were found, of which two-thirds were obtained from Icelandic sites. To reduce the weight of Icelandic observations, only Icelandic records of reversals or excursions yielding both northern and southern VGP latitudes were selected. The final population consists of 401 intermediate VGPs reasonably distributed over the time interval sampled (Fig. 1a). However, both the latitudinal and longitudinal distributions of sites are irregular (Fig. 1b, c), as almost 50% of the data correspond to samples of the Icelandic field.

When calculated using the palaeomagnetic convention (Fig. 2a, b), the longitude distribution of all VGPs does not exhibit maxima in the two opposed longitudinal sectors distinguished by Laj *et al.*³. None of the four statistical tests rejects uniformity when VGPs are all given the same weight (Fig. 2a). Hoffmann's contention¹⁴ that two recurrent long-lived dipolar states dominate the reversal process is based on a selection of VGPs

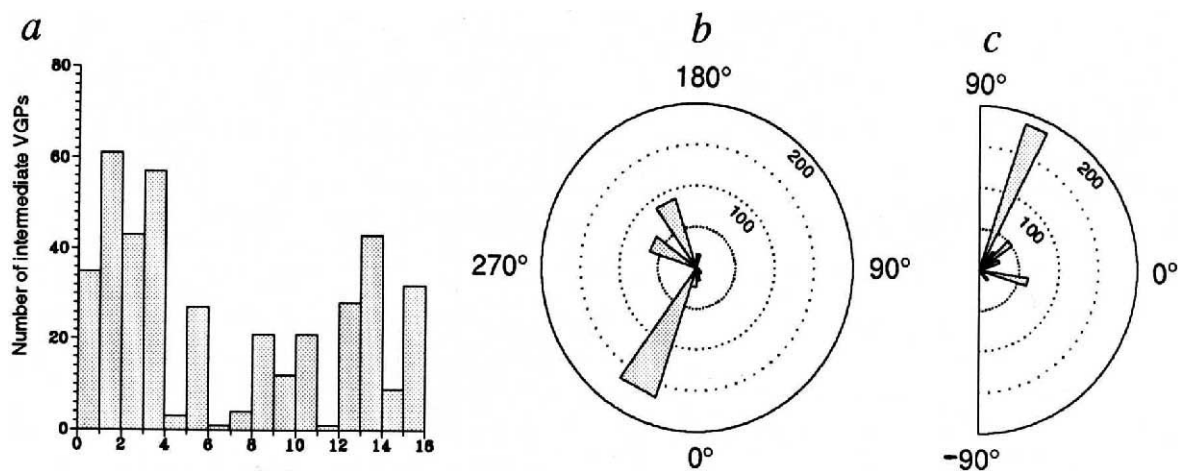


FIG. 1 Characteristics of the database: distribution of age (a), palaeo-longitude (b) and palaeolatitude (c) of the lava flows (or group of consecutive lava flows) having recorded an intermediate virtual geomagnetic

pole (VGP). The length of each column or sector indicates the number of VGPs in this interval (the numbers by the concentric circles refer to absolute numbers of poles).

recorded by at least five consecutive flows. He identified these dipole axes from two VGP clusters centred near the southern tip of South America (mean longitude 315°) and the North West Cape of Australia (mean longitude 110°). Note that this interpretation contradicts the indifference of geomagnetic symmetry axes to the field polarity^{19,20}: two antipodal VGP clusters should be observed for each axis. For comparison, we constructed a variant of our database by giving each VGP a weight equal to the number n of consecutive flows (in a single section) providing this pole (Fig. 2b). In contrast to Hoffmann's result¹⁴, we observe

no obvious clusters of the $n \geq 5$ VGPs on the map. The statistical tests of the longitude distribution of VGPs either accept (Hodges-Ajnes and Rayleigh) or reject (Kuiper and Watson's) uniformity. However the alternative axial Von Mises distribution is extremely scattered ($k=0.06$) and the axis longitude ($32 \pm 180^\circ$) is approximately 90° away from Hoffmann's¹⁴ preferred longitudes. As we will see below, the non-uniformity is due to excursions.

Using the geomagnetic convention reduces the number of VGPs to 362 because the initial polarity of the field before the

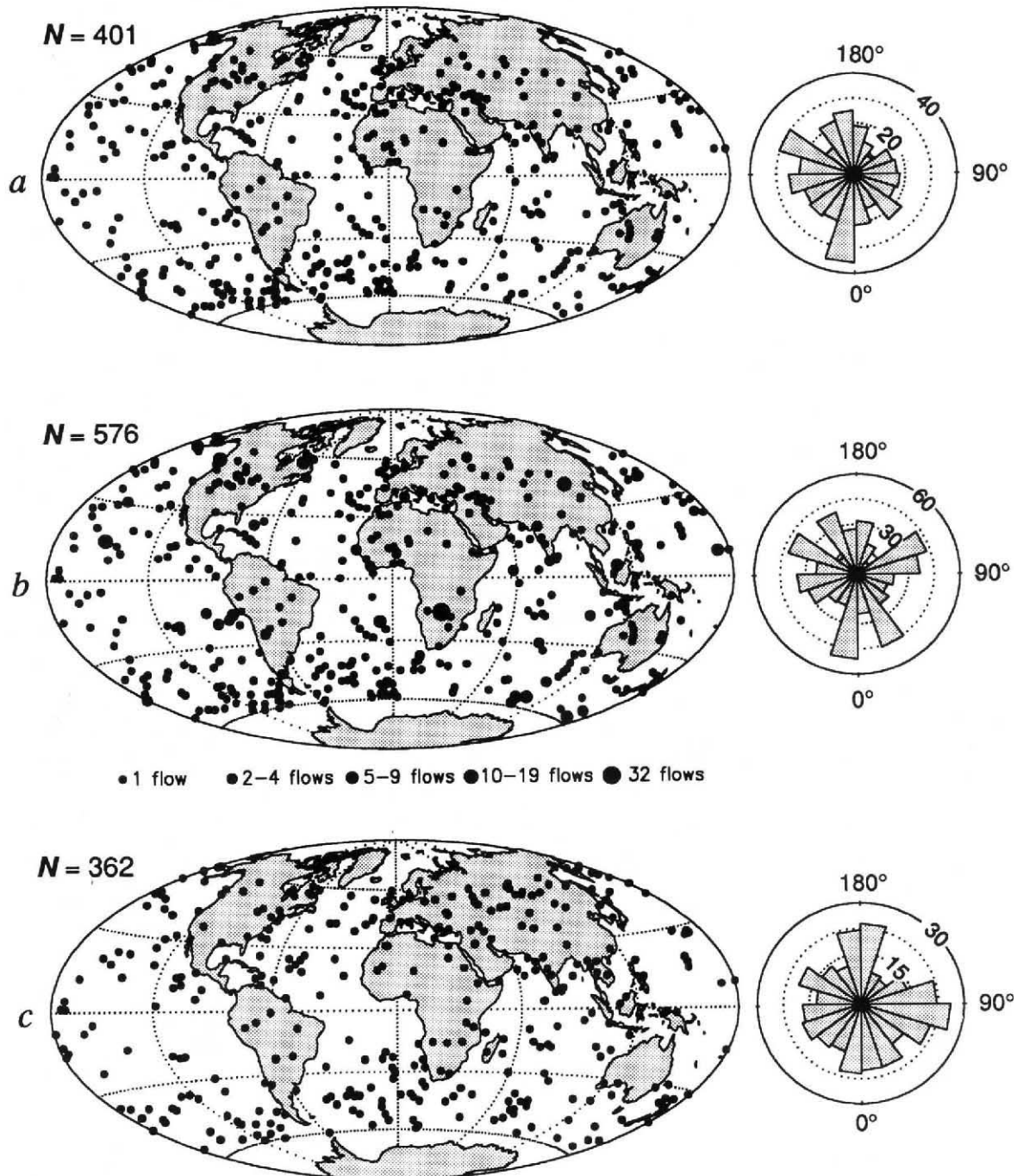


FIG. 2 Geographic location (left) and corresponding circular diagrams of longitude distribution (right) of intermediate VGPs from both excursions and reversals calculated using either the paleomagnetic (a, b) or the geomagnetic (c) conventions (see text). In a (and c), the same weight (one) is used for all VGPs whereas in b weights are equal to the number n of consecutive lava flows from a single section providing 'similar'

VGPs (as defined in the text). Preferred longitude sectors of sedimentary VGPs are centred on 120° and 300° according to Laj *et al.*³ and Hoffmann's¹⁴ VGP clusters are centred on the southern tip of South America and the North West Cape of Australia. None of these maps of geographic location or circular diagrams indicates any obvious VGP cluster or preferred longitude sector.

event has then to be known. The longitude distribution of poles appears uniform (Fig. 2c), which is confirmed by all tests except Watson's. Splitting the VGPs into excursions and reversals shows that the failure to pass Watson's test is due to excursions. This population (Fig. 3) is found not to be uniformly distributed in longitude according to all tests. It is not clear whether this distribution is axial, as suggested by the Icelandic data, or uni-modal, as the non-Icelandic data seem to indicate. More data

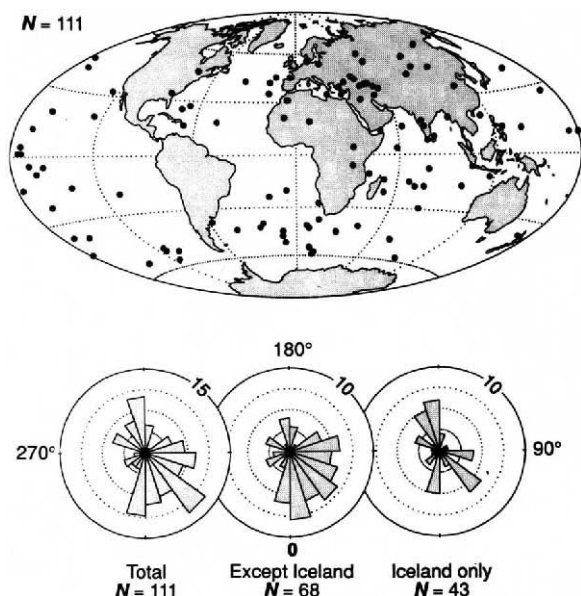


FIG. 3 Geographic location (top) and longitude distribution (bottom) of intermediate VGPs from geomagnetic excursions (see text for precise definition) calculated using the geomagnetic convention. The three diagrams of longitude distribution show that the total population, as well as the Icelandic and the non-Icelandic subsets, exhibit some preference about an axis of longitude close to $30 \pm 180^\circ$.

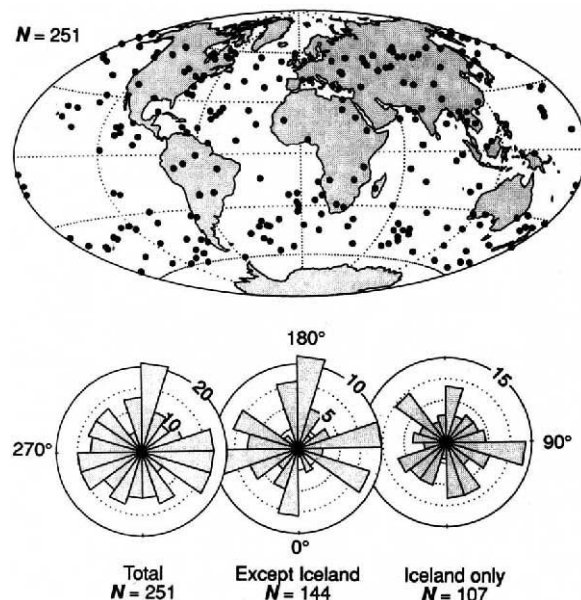


FIG. 4 Geographic location (top) and longitude distribution (bottom) of intermediate VGPs from geomagnetic reversals calculated using the geomagnetic convention. The entire population, as well as the Icelandic and the non-Icelandic subsets, are uniformly distributed in longitude (risk of error 0.05). Thus field reversals are statistically axisymmetrical regardless of the site latitude of observation.

are needed to resolve this question. The overall distribution is approximated by a Von Mises distribution with axis longitude $31 \pm 180^\circ$ and $k = 0.51$. Interestingly, this axis lies almost exactly 90° away from the preferred longitude sectors of the sedimentary VGPs.

In contrast, reversals provide VGPs which are uniformly distributed in longitude (Fig. 4), a conclusion which is confirmed by all tests and applies to the total population ($N = 251$), as well as the Icelandic and non-Icelandic subsets. This conclusion also holds true when the entire population is subdivided into three subsets with VGP latitudes lying in the ranges $0-20^\circ$, $20-40^\circ$ and $40-60^\circ$. This last finding indicates that the absence of longitudinal preference of intermediate VGPs recorded by volcanic rocks during reversals is not an artefact due to the spoiling of the weakened core field by stray fields of crustal origin. This is because it has been convincingly demonstrated^{12,23,24} that the core field intensity decreases with latitude during reversals, whatever the site latitude. Thus, if the core field yields longitudinally organized VGPs whereas the crustal field provides randomly distributed VGPs, the longitudinal organization should become more and more apparent as we consider in our database successive subsets of VGPs with larger latitude. As pointed out above, this is not the case: the distribution remains uniform even for the subset with the highest latitude.

Four main conclusions may be drawn from these results. First, the discrepancy between volcanic and sedimentary records of field reversals points again to the need for a critical re-evaluation of the magnetization blocking process and early remagnetization effects in sediments, as recently attempted by several authors^{5,6,25}.

Second, during reversals the geomagnetic field is statistically axisymmetrical when averaged over time spans of the order of 10^7 yr (and possibly less). There is no evidence of the recently postulated¹⁴ long-lived recurrent inclined dipole states. Thus reversals can be simulated by simply allowing the axial part of the dipole to decay and grow back with opposite sign while the other terms fluctuate about a zero mean as in Constable's model²⁶.

Third, given the absence of a long-term longitudinal structure of the reversing field, there is presently no palaeomagnetic evidence for a control of reversal morphology by the lowest region of the mantle. Our results do not preclude some confinement of the VGP path of some individual reversals—such as, for example, the Matuyama–Brunhes, for which the ten volcanic VGPs so far available^{24,27} fall within one of the two (wide) preferred longitudinal sectors of sedimentary VGPs²⁸. However, these preferential longitude sectors must have changed frequently over the past 16 Myr, to produce an overall uniform distribution. This requires short time constants, which are not compatible with a control of the transitional field by the presumably long-term thermal structure of the lowest region of the mantle.

Finally, the palaeomagnetic data presently available from volcanic rocks indicate that the field behaves similarly at high (Iceland) and lower latitudes. In contrast, the field symmetry seems to differ during reversals, excursions and palaeosecular variation. Thus, it may not be reasonable to try to derive a reversal model from observations of secular variation, as attempted recently^{13,14,19}. □

Received 15 January; accepted 1 September 1993.

1. Laj, C., Mazaud, A., Weeks, R., Fuller, M. & Herrero-Bervera, E. *Nature* **351**, 447 (1991).
2. Tric, E. et al. *Phys. Earth planet. Inter.* **65**, 319–336 (1991).
3. Laj, C., Mazaud, A., Weeks, R., Fuller, M. & Herrero-Bervera, E. *Geophys. Res. Lett.* **19**, 2003–2006 (1992).
4. Rochette, P. *Earth planet. Sci. Lett.* **98**, 33–39 (1990).
5. Van Hoof, A. A. M. & Langeris, C. G. *Nature* **351**, 223–225 (1991).
6. Langeris, C. G., Van Hoof, A. A. M. & Rochette, P. *Nature* **358**, 226–230 (1992).
7. McFadden, P. L., Barton, C. E. & Merrill, R. T. *Nature* **361**, 342–344 (1993).
8. Prévot, M., Mankinen, E. A., Grommé, C. S. & Coe, R. S. *Nature* **316**, 230–234 (1985).
9. Tanguy, J. C. *Geophys. J. Int.* **103**, 281–283 (1990).
10. Coe, R. S. *Geophys. J. R. astr. Soc.* **56**, 369–383 (1979).
11. Dagley, P. & Lawley, E. *Geophys. J. R. astr. Soc.* **36**, 577–598 (1974).

12. Kristjánsson, L. & McDougall, I. *Geophys. J. R. astr. Soc.* **68**, 273–294 (1982).
13. Constable, C. *Nature* **358**, 230–233 (1992).
14. Hoffman, K. A. *Nature* **359**, 789–794 (1992).
15. Valet, J. P., Tucholka, P., Courtillot, V. & Meynadier, L. *Nature* **356**, 400–407 (1992).
16. Mankinen, E. A., Prévot, M., Grommé, C. S. & Coe, R. S. *J. geophys. Res.* **90**, 10393–10416 (1985).
17. Roperch, P. & Duncan, R. A. *J. geophys. Res.* **95**, 2713–2726 (1990).
18. Kristjánsson, L. & Sigurgeirsson, M. *J. Geomagn. Geoelect.* **45**, 275–288 (1993).
19. Gubbins, D. & Coe, R. S. *Nature* **362**, 51–53 (1993).
20. Hoffman, K. A. *Nature* **309**, 90–91 (1984).
21. Mardia, K. V. *Statistics of Directional Data* (Academic, London, 1972).
22. Jones, G. M. *J. geophys. Res.* **82**, 1703–1709 (1977).
23. Prévot, M., Mankinen, E. A., Coe, R. S. & Grommé, C. S. *J. geophys. Res.* **90**, 10417–10448 (1985).
24. Chauvin, A., Roperch, P. & Duncan, R. A. *J. geophys. Res.* **95**, 2727–2752 (1990).
25. Quidelleur, X., Valet, J. P. & Thouveny, N. *Earth planet. Sci. Lett.* **111**, 23–39 (1992).
26. Constable, C. *J. geophys. Res.* **95**, 4587–4596 (1990).
27. Kristjánsson, L., Johannesson, H., Eriksson, J. & Gudmundsson, A. I. *Can. J. Earth Sci.* **25**, 215–225 (1988).
28. Clement, B. M. *Earth planet. Sci. Lett.* **104**, 48–58 (1991).

SUPPLEMENTARY INFORMATION. Requests should be addressed to Mary Sheehan at the London editorial office of *Nature*.

ACKNOWLEDGEMENTS. Work supported by the French INSU-DBT program.

Oldest known amphisbaenian from the Upper Cretaceous of Chinese Inner Mongolia

Xiao-chun Wu^{*†‡§}, Donald B. Brinkman^{*†},
Anthony P. Russell^{*†}, Zhi-ming Dong[†],
Phillip J. Currie^{*†}, Lian-hai Hou[†]
& Gui-hai Cui[†]

^{*} Department of Biological Sciences, The University of Calgary,
2500 University Drive NW, Calgary T2N 1N4, Alberta, Canada

[†] Royal Tyrrell Museum of Palaeontology, Box 7500,
Drumheller TOJ 0Y0, Alberta, Canada

[‡] Institute of Vertebrate Palaeontology and Paleoanthropology,
Academia Sinica, Box 643, Beijing 100044, China

[§] To whom correspondence should be addressed at the Royal Tyrrell
Museum of Pathology.

THE amphisbaenians, lizards and snakes constitute a monophyletic group, the Squamata. Although amphisbaenians are known to have occurred in the Mesozoic^{1,2}, their remains are rare and fragmentary. The oldest and most primitive known skull of an amphisbaenian is from the Eocene of North America³ and differs little from

that of modern taxa^{4–9}. A substantial structural gap exists between this skull and that of the several possible sister groups^{8,10}. No derived features uniquely linking amphisbaenians to any other group of the Squamata have been recognized, so the relationship of amphisbaenians is uncertain^{11,12}. We report here on the late Cretaceous lizard-like amphisbaenian represented by well-preserved cranial and postcranial material from the Bayan Mandahu redbeds of the Gobi Desert, Inner Mongolia, China. This material documents the oldest and most primitive amphisbaenian yet known, and permits a re-evaluation of the relationship between the amphisbaenians and other squamates.

Squamata

Amphisbaenia or Annulata

Sineoamphisbaena hexatabularis gen. et sp. nov.

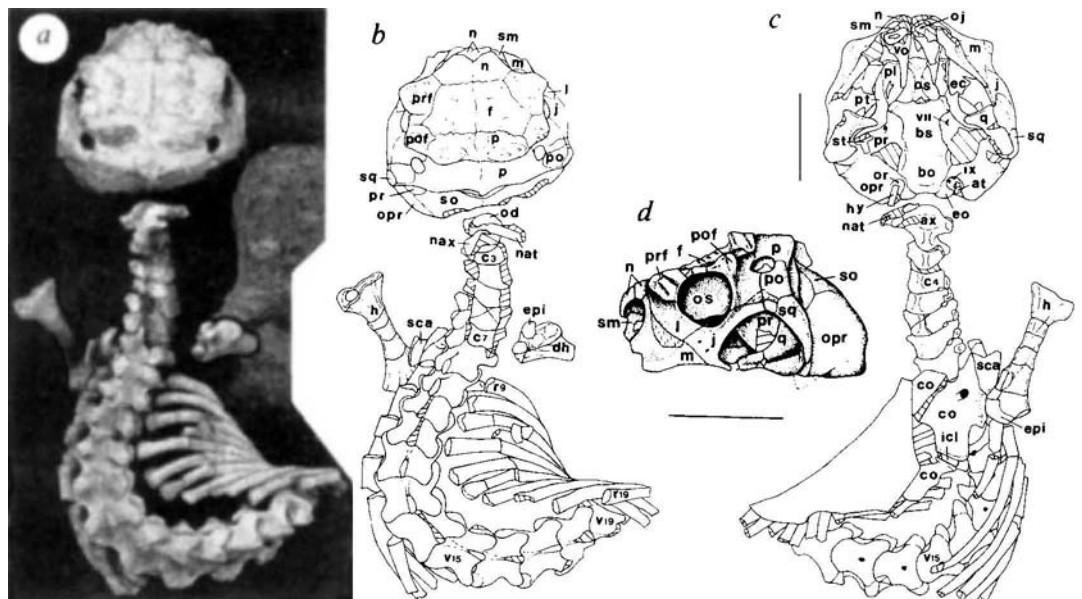
Etymology. The generic name refers to China and to the primitive, less specialized morphology of the new species; the specific name refers to the hexagonal interorbital skull table.

Holotype. IVPP V10593-1, a skull with partial skeleton (Fig. 1).

Paratype. IVPP V10593-2, a skull with left mandible, its roof elements heavily damaged by erosion and its palate not preserved (Fig. 2a).

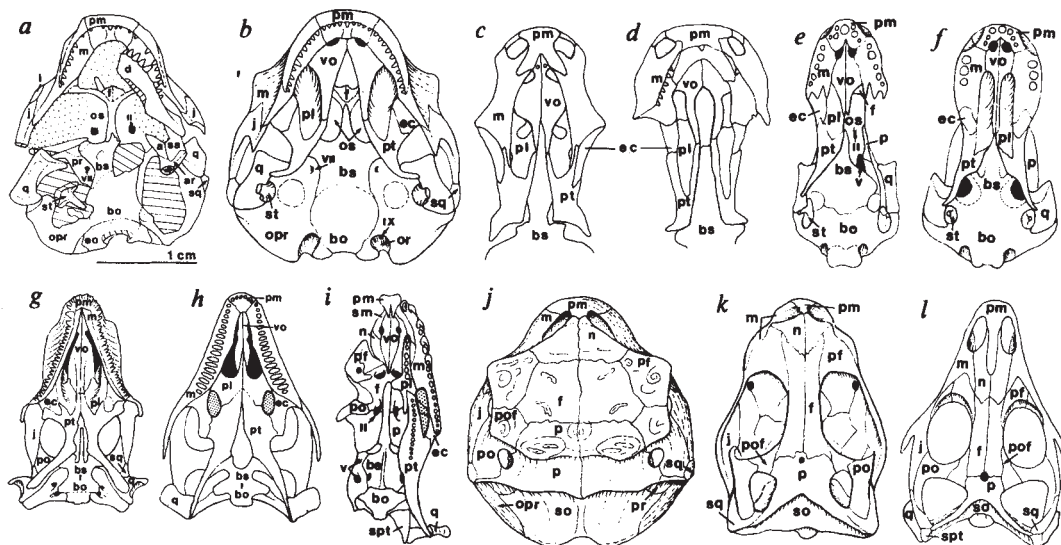
Type locality and horizon. Bayan Mandahu, Inner Mongolia, China. The Bayan Mandahu redbeds correlate with the Upper

FIG. 1 *Sineoamphisbaena hexatabularis* gen. et sp. nov., holotype (IVPP V10593-1) from the Upper Cretaceous of Inner Mongolia, China. Skeleton in dorsal (a), ventral (b) and lateral (c) views; skull in lateral (d) view. Scale bars, 1 cm. Abbreviations: at, part of neural arch of atlas; ax, axis; bo, basioccipital; bs, basisphenoid; c3, c4, c7, cervical vertebrae 3, 4, 7; co, coracoid; dh, distal part of right humerus; ec, ectopterygoid; eo, exoccipital; epi, epiphysis; f, frontal; h, humerus; h, part of hyoid; icl, part of interclavicle; j, jugal; l, lacrimal portion of fused prefrontal/lacrimal complex; m, maxilla; n, nasal; nat, neural arch of atlas; nax, neural arch of axis; od, odontoid process; oj, opening for Jacobson's organ; opr, opisthotic; or, occipital recess; os, orbitosphenoid; p, parietal; pl, palatine; po, postorbital portion of fused postfrontal/postorbital complex; pof, postfrontal portion of fused postfrontal/postorbital complex; pr, prootic; prf, prefrontal portion of fused prefrontal/lacrimal complex; pt, pterygoid; q, quadrate; r9, r19, ribs 9, 19; sca, scapula; sm, septomaxilla; so, supraoccipital; sq, squamosal; vo, vomer; v15, v19, vertebrae 15, 19; VII, IX, foramina for cranial nerves. The hexagonal interorbital region of *S. hexatabularis* is crowned by elevated dermal rugosities that form a distinct table contributed to by five pairs of roof elements. Some chamaeleons, such as *Rhampholeon spectrum*²³, have a hexagonal



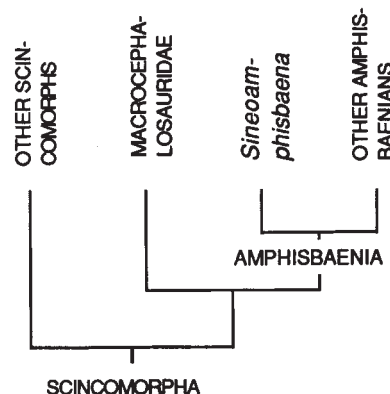
interorbital region, but it consists only of an unpaired frontal and does not form a table. The medial process of the palatal ramus of the pterygoid is believed to have been broken away in *S. hexatabularis*. Although the suture between the orbitosphenoid and frontal is obscured by fusion, the presence of the orbitosphenoid is indicated by its contacts with the prootic posteriorly, its contribution to the border of the trigeminal foramen, and the fact that foramina for the optic nerves are located entirely within this bone.

FIG. 2 Skulls and palates in ventral (a–f) and dorsal (j–l) views. a, b, j, *Sineoamphisbaena hexatabularis*: a, paratype, IVPP V10593-2; b, j, reconstruction of holotype, IVPP V10593-1; c, Middle Eocene amphisbaenian *Spathorhynchus fossorium*³; d, oligocene amphisbaenian *Hyporhina galbreathi*⁶; e, f, two living amphisbaenians: e, *Amphisbaena alba* (no. 54299, Museum of Comparative Zoology-MCZ); f, *Bipes biporus* (no. 145823, MCZ); g, k, two macrocephalosaurid lizards²¹: g, *Macrocephalosaurus gilmorei*; k, *Darchansaurus estesi*; h, l, a polyglyphanodontid lizard^{21,22}; i, a living snake *Python regius*²⁴. Illustrations unscaled except for a. Abbreviations as in Fig. 1 legend plus a, angular; ar, articular; d, dentary; pm, premaxilla; sa, surangular; spt, supratemporal; II, V, foramina for cranial nerves. The anterior closure of the braincase was thought to be an amphisbaenian synapomorphy¹⁰, but others argued that it was a derived feature shared by amphisbaenians and snakes¹¹. We agree with the former conclusion¹⁰ in believing that these two patterns differ. The parabasisphenoid process meets the frontal dorsolaterally and floors the braincase in snakes, but the descending processes of the frontals rest on the orbitosphenoid and the parabasisphenoid process



underlies this in amphisbaenians. The optic nerves penetrate only the orbitosphenoid in living amphisbaenians²⁵; this feature is considered here to be unique to the Amphisbaenia. In *S. hexatabularis*, other fossil species that have the pterygoid preserved, and all living species of the Amphisbaenia, the quadrate ramus of the bone is very short and is tightly wrapped around the quadrate (the latter situation can not be tested in other fossil forms because of the occlusion of the lower jaws or poor preservation). This is not the case in lizards and snakes.

FIG. 3 A cladogram depicting the sister group relationship of the Amphisbaenia to macrocephalosaurid lizards as suggested in the text. This pattern was obtained by using PAUP to analyse a character matrix¹¹ and including relevant fossil lizards; details of this analysis are to be presented elsewhere. In certain living agamid and lacertid lizards the medial process of the palatal ramus of the pterygoid meets, or nearly meets, the vomer, and in certain scincomorph lizards, the lateral process of the palatal ramus contacts the maxilla. But no squamates other than amphisbaenians and macrocephalosaurids have a forked palatal ramus of the pterygoid that simultaneously contracts or approaches the vomer medially, meets the ectopterygoid (sometimes the maxilla too) laterally, and closes the suborbital fenestra. The separation of the postorbital from the infratemporal fenestra, because of a contact between the squamosal and jugal, occurs in several lizard groups. However, the postorbital in these lizards is relatively small, posteriorly tapered and widely separated from the parietal. More derived amphisbaenians have all lost the temporal fenestrae and most of the associated elements.



Cretaceous Djadokhta Formation of Mongolia (no younger than Campanian)¹³.

Diagnosis. Lizard-like amphisbaenian, differing from all other squamates in possessing a hexagonal interorbital table with a pair of pronounced parietal rugosities; postfrontal/postorbital complex contacts prefrontal/lacrimal complex above the orbit, and parietal both anterior and posterior to supratemporal fenestra; small supratemporal fenestra enclosed only by parietal and postfrontal/postorbital complex; slender supratemporal process of parietal directed ventrally and anterolaterally; dermal skull roof wider than long; squamosals and parietals form semicircular arc tightly embracing braincase in occipital view; and each anterior dorsal vertebra bears single median foramen on ventral midline of centrum.

Sineoamphisbaena hexatabularis is referred to the Amphisbaenia on the basis of the following unique synapomorphies: (1) skull solidly built^{10,11}; (2) enlarged median tooth on fused premaxillae^{10,11}; (3) unpaired dermal orbitosphenoid^{10,11}; (4) anterior end of braincase closed by single orbitosphenoid;

descending processes of frontals rest on the orbitosphenoid¹⁰; (5) pair of tiny optic foramina situated entirely within the orbitosphenoid; (6) quadrate ramus of pterygoid very short and wraps around posteromedial (ventromedial if quadrate horizontally oriented) surface of quadrate (Figs 1, 2a–f, j).

In addition, *S. hexatabularis* shares with the Amphisbaenia a number of putative synapomorphies¹¹, although these are not unique to the group: (1) prefrontal/postfrontal contact above the orbit; (2) descending processes of frontals in contact below olfactory tracts; (3) supratemporal lost; (4) epipterygoid lost; (5) large, lateroventrally placed fenestra ovalis; (6) subdental shelf small or lost; and (7) pterygoid teeth lost.

S. hexatabularis is the most primitive known species within the Amphisbaenia, retaining numerous lizard-like (primitive) characters: (1) supratemporal fenestra present; (2) orbit relatively large; (3) nares not terminal or anteroventral in position; (4) parietals separate; (5) prefrontal contacts nasal; (6) postorbital present (but fused to postfrontal); (7) anterior skull elements not strongly interdigitating; (8) pronounced posterior

process of jugal present; (9) exoccipital not fully fused with opisthotic; (10) lateral aperture for recessus scala tympani open; (11) interclavicle present; (12) ectepicondylar foramen present on humerus.

The primitive morphology of *S. hexatabularis* within the Amphisbaenia allows a reevaluation of the hypotheses of relationships of this group within the Squamata. Previous studies^{11,14-16} have agreed that the sister group of the Amphisbaenia lies within the Scleroglossa, which includes all squamate families other than those of the Iguania¹¹. Many of the studies suggest a close relationship between the Amphisbaenia and scincomorph lizards^{11,17-20}, but there has been no consensus. We propose a relationship within the scincomorpha between the Amphisbaenia and the Macrocephalosauridae²¹ (part of the Polyglyphanodontinae²²) on the basis of uniquely shared, derived features of the palate and temporal region (Fig. 3): the palatal ramus of the pterygoid is forked into a medial process, which extends along the medial side of the palatine and contacts or nearly reaches the vomer, and a lateral process which meets the ectopterygoid (sometimes the maxilla too) along the lateral side of the palatine and closes the suborbital fenestra (Fig. 2b-g); and a large postorbital is posteriorly broad, extends posteriorly over the temporal fenestrae, meets or closely approaches the parietal, and is excluded from the infratemporal fenestra by the contact of the jugal with the squamosal (Figs 2j, k; 4a). In addition, macrocephalosaurids occur in the same area of central Asia and the same geological period as *S. hexatabularis*.

Despite its well developed limbs and less specialized skull than that of later forms (Fig. 4b-d), the tiny supratemporal fenestra, heavily rugose dermal elements of the skull roof, solidly built cranium, posterolaterally facing orbit, large otic capsule with reduced occipital recess, and the massive stapes strongly suggest that *S. hexatabularis* was adapted to a fossorial way of life. □

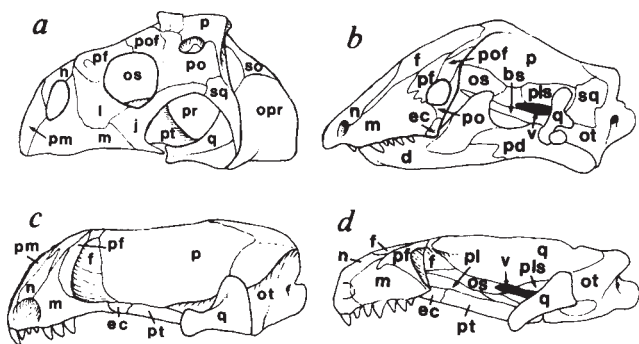


FIG. 4 Lateral views of four amphisbaenian skulls, showing a sequence of reductions of the supratemporal fenestral and orbit, and specializations of the skull for tunnelling. a, Reconstruction of holotype (IVPP V10593-1) of *Sineoamphisbaena hexatabularis*; b, a Middle Eocene amphisbaenian *Spathorhynchus fossorium*³; c, a living amphisbaenian *Bipes biporus* (the only living species with functional forelimbs), no. 145823, Museum of Comparative Zoology (MCZ); d, a living amphisbaenian *Amphisbaena alba*, no. 54299, MCZ. Illustrations are unsealed. Abbreviations as in previous figures, plus ot, fused otic/occipital complex; pd, fused postdentary complex; pls, pleurophenoid (which is totally co-ossified with the prootic⁶). The skull of *S. hexatabularis* seems to be less specialized for tunnelling in soils when compared to those of later fossil and living species, which have an elongate, low skull with a spade-shaped or keeled snout. These differences may mean that *S. hexatabularis* lived in an environment very different from that of later fossil forms or living species. The Bayan Mandahu locality is thought to have been situated at the southern margin of an extensive sand sea dominated by eolian dunes²⁶. This site is characterized by structureless eolian deposits with minor lacustrine and interdune fluvial deposits. *S. hexatabularis* may have tunnelled in the eolian deposits, using its limbs to dig tunnels and its broad heavily armoured interorbital skull table to ram-loose tunnel walls.

Received 28 April; accepted 25 August 1993.

1. Borsuk-Bialynicka, M. *Cretac. Res.* **12**, 607-608 (1991).
2. Astibia, H. et al. *Terra Res.* **2**, 460-466 (1991).
3. Berman, D. *Copeia* **4**, 704-721 (1973).
4. Gilmore, C. *Mem. natn. Acad. Sci.* **22**, 1-197 (1928).
5. Taylor, E. H. *Kans. Univ. Sci. Bull.* **34**, 521-579 (1951).
6. Berman, D. J. *Paleont.* **50**, 165-174 (1976).
7. Berman, D. J. *Paleont.* **51**, 986-991 (1977).
8. Zangerl, R. *Am. Midl. Nat.* **33**, 764-780 (1945).
9. Gans, C. *Bull. Am. Mus. Nat. Hist.* **119**, 129-204 (1960).
10. Gans, C. *Trans. zool. Soc. Lond.* **34**, 347-416 (1978).
11. Estes, R., Queiroz, K. & Gauthier, J. in *Phylogenetic Relationships of the Lizard Families* (eds Estes, R. & Pregill, G.) 120-280 (Stanford Univ. Press, California, 1988).
12. Rieppel, O. in *The Phylogeny and Classification of the Tetrapods Vol. 1* (ed. Benton, M. J.) 261-294 (Clarendon, Oxford, 1988).
13. Jerzykiewicz, T. et al. *Can. J. Earth Sci.* (in the press).
14. Rage, J.-C. *Pour La Sci.* **1982**, 16-27 (1982).
15. Greer, A. E. J. *Herpetol.* **19**(1), 116-156 (1985).
16. Presch, W. in *Phylogenetic Relationships of the Lizard Families* (eds Estes, R. & Pregill, G.) 471-492 (Stanford Univ. Press, California, 1988).
17. Bogert, C. *Nat. Hist.* **73**, 17-24 (1964).
18. Rieppel, O. *J. Zool.* **195**, 493-528 (1981).
19. Böhme, W. in *Handbuch der Reptilien und Amphibien Europas Vol. 1* (ed. Böhme, W.) 275-276 (Akademische Verlagsgesellschaft, Wiesbaden, 1981).
20. Schwenk, K. in *Phylogenetic Relationships of the Lizard Families* (eds Estes, R. & Pregill, G.) 569-598 (Stanford Univ. Press, California, 1988).
21. Sulimski, A. *Palaeont. pol.* **33**, 25-102 (1975).
22. Estes, R. *Handbuch der Paläoherpetologie Vol. 10A* (Fischer, Stuttgart, 1983).
23. Rieppel, O. *Zool. J. Linn. Soc.* **89**, 41-62 (1987).
24. Rage, J.-C. *Handbuch der Paläoherpetologie Vol. 11* (Fischer, Stuttgart, 1984).
25. Bellairs, A. D. A. & Gans, C. *Nature* **302**, 243-244 (1983).
26. Eberth, D. A. *Can. J. Earth Sci.* (in the press).

ACKNOWLEDGEMENTS. The specimens of *S. hexatabularis* were collected during the Dinosaur Project, an exchange between the Institute of Vertebrate Paleontology and Paleoanthropology (IVPP), the Canadian Museum of Nature, Ottawa, and the Royal Tyrrell Museum of Paleontology (RTMP). They are housed in the IVPP. X.-C.W. thanks IVPP for allowing him to study the material and the RTMP for access to facilities. We thank D. A. Eberth for discussion, J. Rosado for the loan of skeletons of extant amphisbaenians, and C. Coy for photographing the type. X.-C. W. holds a University of Calgary postdoctoral fellowship; A.P.R. is supported by an NSERC operating grant.

An identified neuron mediates the unconditioned stimulus in associative olfactory learning in honeybees

Martin Hammer

Institut für Neurobiologie, Freie Universität Berlin, Königin-Luise Strasse 28/30, 14195 Berlin, Germany

DURING classical conditioning, animals learn to associate a neutral stimulus with a meaningful, or unconditioned, stimulus. The unconditioned stimulus is essential for forming associations, and modifications in the processing of the unconditioned stimulus are thought to underlie more complex learning forms¹⁻⁴. Information on the neuronal representation of the unconditioned stimulus is therefore required for understanding both basic and higher-order features of conditioning. In honeybees, conditioning of the proboscis extension reflex occurs after a single pairing of an odour (conditioned stimulus) with food (unconditioned stimulus)^{5,6} and shows several higher-order features of conditioning⁶⁻⁸. I report here the identification of an interneuron that mediates the unconditioned stimulus in this associative learning. Its physiology is also compatible with a function in complex forms of associative learning. This neuron provides the first direct access to the cellular mechanisms underlying the reinforcing properties of the unconditioned stimulus pathway.

Neurons mediating the reinforcing property of the unconditioned stimulus (US) in conditioning must meet two criteria: response to the US and convergence with the conditioned stimulus (CS) pathway. To identify such neurons in conditioning of the proboscis extension reflex, an *in vivo* preparation was developed in which single cells in the suboesophageal ganglion were impaled, tested for their responsiveness to the US (sucrose solu-

tion applied briefly to the antennae and proboscis), and subsequently stained with Lucifer yellow. One particular candidate neuron, VUMmx1, was identified by its specific response characteristics and unique morphology (Fig. 1). It responds to sucrose with a long burst of action potentials (duration ~ 30 s; frequency over 30 s, 4.6 ± 0.5 Hz (mean, s.e.m.); $n = 10$) which outlasts the US. VUMmx1 arborizes in the dorsal suboesophageal ganglion and innervates brain neuropiles involved in odour (CS) processing: that is, the antennal lobe glomeruli, the lateral protocerebrum and the mushroom body calyces^{9,10}. It therefore shows multiple convergence onto CS-processing neuropiles.

As an appetitive stimulus, the US has several properties. It releases behavioural acts, such as the proboscis extension reflex, transiently sensitizes feeding behaviour (non-associative modulations)¹¹, and serves as a reinforcer in conditioning. As these properties may be processed in parallel at the neural level, all neurons responding to sucrose do not necessarily mediate conditioning. To investigate whether VUMmx1 meets a third essential criterion for mediating reinforcement in learning, bees were conditioned to an odour in a single trial, in which sucrose as a US was replaced by a supra-threshold depolarization of VUMmx1 (Fig. 2a).

Bees only learn to respond to an odour with the proboscis extension reflex (conditioned response) when it is presented shortly before the US during conditioning⁵ (Fig. 2c, open bars); that is, associative learning depends on temporal contiguity between CS and US. Thus, by presenting an odour either before (forward pairing; $n = 12$) or after (backward pairing; $n = 11$) the start of depolarizing VUMmx1 (Fig. 2a), associative effects could be discriminated from non-associative effects. Odour responses were quantified by comparing odour-evoked electromyograms from muscle M17, which controls the proboscis extension reflex¹², in 10-s intervals after odour onset 5 min before (pre-test) and 10 min after (test) pairing (Fig. 2b). For forward pairing, the number of M17 spikes increased from pre-test to test (Fig. 2c) (pre-test, 1.5 (0,11); test, 61 (28,101.5); median of spikes (interquartile ranges); $P < 0.05$). Occasionally M17 activity increased during forward pairing (median: 42.5 (1.5,61.5) spikes, NS), which might indicate that coincidence of an odour and VUMmx1 activity immediately alters odour processing. VUMmx1 probably does not directly contribute to PER, however, because its depolarization does not evoke motor activity during backward pairing. Furthermore, the backward pairing group reveals that replacing a sucrose reward by depolarizing

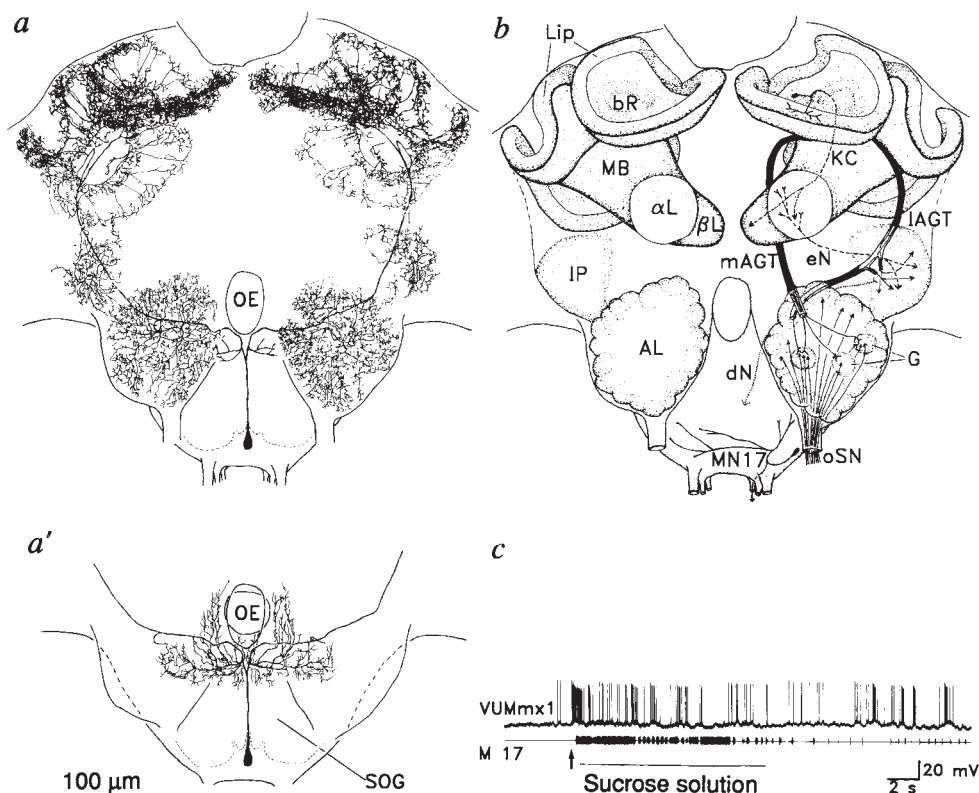


FIG. 1 Morphology of VUMmx1 (ventral unpaired median cell of maxillary neuromere 1) and its response to sucrose (US). *a* and *a'*, Morphology: *a*, VUMmx1 innervates the glomeruli (G in *b*) of the antennal lobes (AL), the lateral protocerebrum (IP), and lips and basal rings (bR) of the mushroom bodies (MBs) calyces (for orientation, see *b*); *a'*, in the suboesophageal ganglion (SOG) the primary neurite projects dorsally from the ventral median soma and bifurcates beyond the oesophagus (OE). Dendritic arborizations occur in the dorsal SOG and the tritocerebrum. VUMmx1 is one of a group of VUM neurons in the maxillary neuromere, but can be reliably identified by its unique morphology. *b*, Scheme of the olfactory pathway. Olfactory sensory neurons (oSN) from the antennae serve the glomeruli. Relay neurons connect glomeruli with the IP and the calyces via two fibre tracts (IAGT and mAGT)^{9,10}. In bees, the MBs calyces receive various modalities separately: the lip primarily olfactory, the collar visual, and the basal ring multimodal, thought to arise partly from collaterals of visual and olfactory fibres¹⁰. Kenyon cells

(KC) distribute information to the MBs output lobes (α L and β L). Some extrinsic neurons (eN) connect the MB with the IP²⁴. Motor neurons (such as MN17) generating the proboscis extension reflex (PER) are located in the ventral SOG²⁵. In insects, descending neurons (dN) are thought to connect the IP with motor centres²⁶. *c*, Long-lasting excitation of VUMmx1 following application of sucrose to the antennae and proboscis (arrow). Little, if any, phasic excitation is elicited by visual and mechanical stimuli to the mouthparts and antennae. Lower trace: corresponding PER recorded as an electromyogram from M17.

METHODS. Bees were fixed by waxing the dorsal thorax to a small metal table, which allowed free movements of mouth parts and antennae. Impaled neurons were stained with Lucifer yellow (5%; hyperpolarizing current pulses, 2–8 nA, for 10–40 min). As in conditioning experiments, sucrose solution ($\sim 30\%$) was briefly (< 1 s) delivered to the antennae and proboscis. Odours (carnation) were presented in a gentle air stream. Stimulus duration was controlled with a solenoid valve.

VUMmx1 during conditioning alters CS processing in a way that is dependent only on a specific associative feature: temporal contiguity between the CS and VUMmx1 activity. The test response after forward pairing depends on the temporal relations during pairing as it significantly exceeds that after backward pairing (Fig. 2c) ($P < 0.002$). This difference cannot be due to a different spike frequency of VUMmx1 during depolarization (forward, 7.8 ± 0.7 Hz; backward, 7.9 ± 0.7 Hz).

Using sucrose as the US under the same experimental conditions, without intracellular recording, the response to an odour after forward pairing (median: 70(21,135) M17 spikes; $n = 15$; Fig. 2c) was no different from that recorded in substitution experiments. Thus, VUMmx1 is an equally effective US as is sucrose. But as VUMmx1's spike frequency in substitution

experiments exceeds its sucrose response, other neurons may also be involved in mediating the US.

The observation that odours elicit a short phasic excitation of VUMmx1 (Figs 2b, 3a) raises the possibility that the efficacy of the olfactory input to VUMmx1 is altered after pairing. Quantitative analysis of VUMmx1's odour response (Fig. 3) reveals an increase in the total number of VUMmx1 spikes from 12.1 ± 2.5 (pre-test) to 26.0 ± 6.4 spikes (test) for forward ($P < 0.025$) but not backward pairing (pre-test, 11.7 ± 2.5 ; test, 9.2 ± 1.7 ; $P < 0.05$ between tests after forward and backward pairing). An analysis of the spike frequency in 15 consecutive 1-s intervals indicates that the increased response after forward pairing is due to a long-lasting excitation (Fig. 3b). Temporal coincidence of activity in the olfactory system and VUMmx1

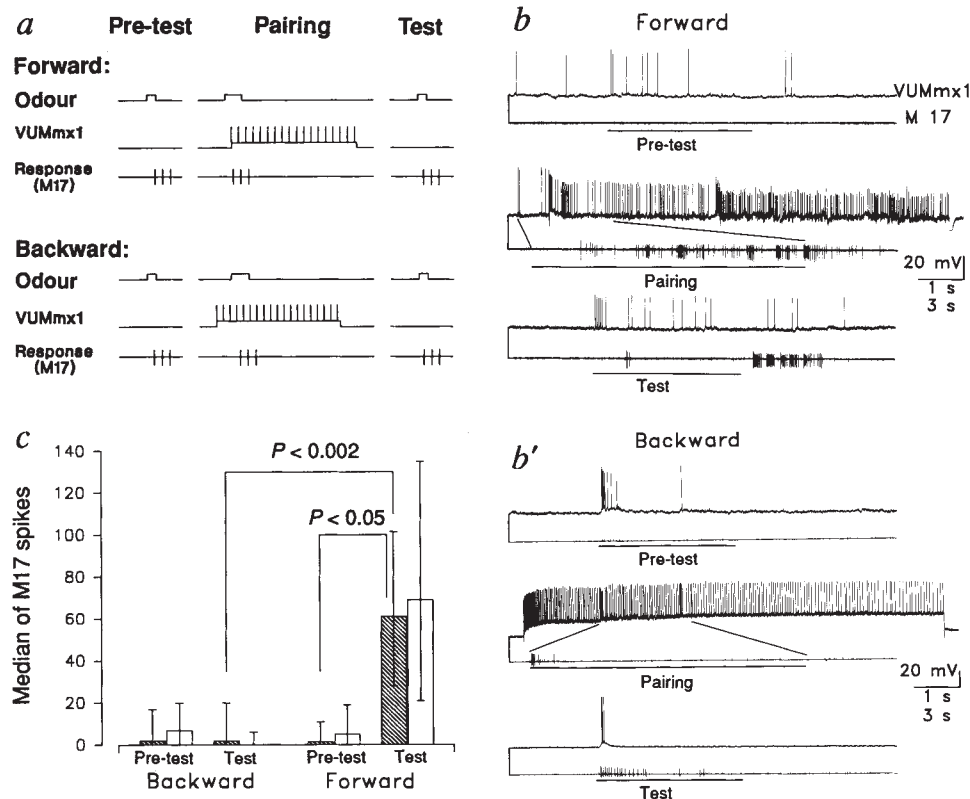


FIG. 2 Substitution of VUMmx1-depolarization for the US in PER conditioning increases subsequent odour responses of bees when the temporal relations between odour stimulus and VUMmx1 activity are correct. **a**, Experimental design. During pairing, an odour (duration, 6 s) is delivered in temporal proximity with spike activity of VUMmx1 evoked by current injection (duration, 30 s). In forward pairing, odour onset precedes depolarization by 2 s; in backward pairing, it follows 5 s after onset of depolarization. Effects of treatment are recorded as electromyograms (EMGs) from M17 by comparing odour responses 5 min before (pre-test, stimulus duration 3 s) and 10 min after pairing (test, duration 3 s). **b** and **b'**, Response change induced by depolarizing VUMmx1. Intracellular recordings from VUMmx1 (upper traces) and EMGs from M17 (lower traces) before, during and after substitution of a sucrose reward by depolarizing VUMmx1. VUMmx1 depolarization is shown on a slower timescale. (Calibration scale: lower time value, upper of middle two traces only; upper value, all other traces.) In **b**, an example is shown of a forward pairing of odour and depolarization (recording trace during depolarization is displayed a.c.-coupled); **b'** shows an example of a backward pairing. **c**, Quantitative comparison of the effect of the substitution experiment (hatched bars; $n = 11$ for backward, $n = 12$ for forward). Data are shown for comparison from a parallel group of control

bees (without intracellular recording) trained with either forward or backward pairing, in which sucrose was used as US (open bars, $n = 15$ for each group). Data are expressed as median (with interquartile ranges) number of M17 spikes in a 10-s interval after odour onset.

METHODS. Timing between CS and VUMmx1 depolarization was adjusted according to behavioural studies⁵. Test odours were presented 10 min after pairing, because retrieval in PER conditioning is exclusively controlled by an associative memory ~ 7 min after a single learning trial^{6,11}. Depolarization current of 0.8–1.6 nA was injected to evoke spike activity of VUMmx1. As sufficient balancing of the microelectrode resistance was not always possible, recordings were stored d.c.- and a.c.-coupled. Impaled cells were identified as VUMmx1 by staining with Lucifer yellow. Bees were stimulated with sucrose at 3–5 min after test to ensure that the failure of the backward-paired bees to show PER was not due to performance impairment (median of M17 spikes: forward, 315 (264,387); backward, 301 (252,430)). Four bees per group showing no or an impaired PER were excluded from analysis. Non-parametric statistics for independent (*U*-test) or correlated data sets (Wilcoxon test), respectively, were used for data sets with cutoffs or without normal distribution. Appropriate parametric statistics were applied in all other cases (*t*-test, paired *t*-test). Two-tailed statistics were used throughout.

thus induces an increase of the olfactory input to VUMmx1 which may depend on the associative strength of odours. Monitoring VUMmx1's activity during differential conditioning supports this hypothesis (Fig. 4). Differentially conditioned bees associate one odour (CS^+) specifically paired with the US as an indicator for food, but do not associate another specifically unpaired odour (CS^-)⁷. After successful differential conditioning, VUMmx1's response to the CS^+ increases compared with pre-test (Fig. 4b; pre-test, 11.4 ± 3.9 ; test, 41.3 ± 10.6 spikes; $P < 0.015$; $n = 5$). This difference is pairing-specific. The test response to the CS^+ significantly exceeds that to the CS^- (7.8 ± 4.0 ; $P < 0.03$). VUMmx1 responds to the CS^+ with a prolonged excitation (Fig. 4b). Thus, VUMmx1 has a discharge to a learned odour similar to that produced by the US (Fig. 1c).

Tracing modulatory or reinforcing pathways has helped to validate ideas about cellular mechanisms of plasticity and to scrutinize sites of plasticity in non-associative¹³ and associative learning models^{14,15}. My study provides evidence that a single neuron is sufficient to mediate reinforcement in associative learning. But the difference in the strength of VUMmx1's sucrose response and the depolarization effective in substitution experiments suggests that other neurons participate in mediating the US. A serial homologue of VUMmx1, for example, has an almost identical branching pattern and also responds to sucrose. Correlating VUMmx1 activity in substitution and conditioning experiments, in which sucrose is used as US, with learning success should provide evidence for this interpretation.

The distribution of branches of VUMmx1 indicates crucial

brain structures for associative learning. VUMmx1 converges with the CS pathway in three different brain regions, reinforcing the concept that integration underlying learning and memory occurs at multiple sites. This and other studies in bees^{16,17} and fruitflies¹⁸⁻²¹ implicate the mushroom bodies in such higher brain functions. The restriction of VUMmx1's branches in the mushroom bodies to the calyces is the first indication that memory induction occurs at the mushroom body input sites. VUMmx1's projections in the antennal lobes and lateral protocerebrum, however, emphasize that these structures may also be involved in associative learning.

VUMmx1 fails to evoke motor activity during backward pairing, indicating parallel processing of reflex and reinforcement. This could allow the uncoupling of reflexive behaviour from alterations in the effectiveness of the reinforcing pathway, which may underlie more complex forms of learning. The plasticity of VUMmx1's odour response suggests its involvement in processing higher-order features of conditioning. After forward pairing, the odour elicits a long-lasting discharge of VUMmx1, indicating that its responsiveness to an odour could depend on the associative strength of that odour. The difference of VUMmx1's odour response to two differentially conditioned odours confirms this hypothesis. Learning therefore adds more to a CS than the capacity to evoke a conditioned response: a learned CS also activates VUMmx1. VUMmx1 can thus maintain its functional property of mediating the reinforcer in the absence of the US, a feature that could underlie second-order conditioning, for instance^{4,22,23}. □

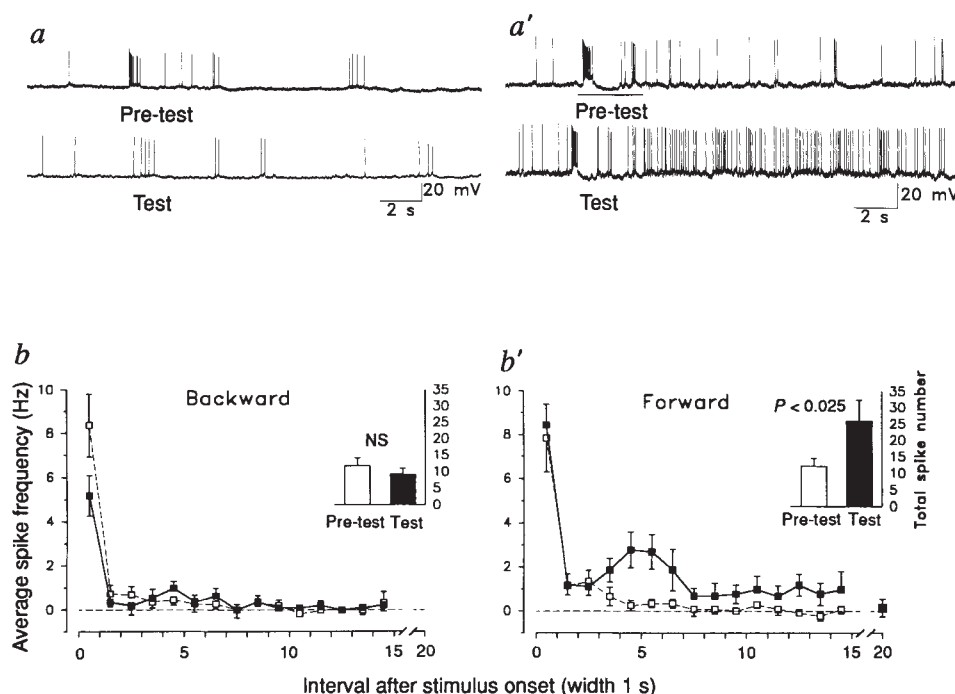


FIG. 3 Plasticity of VUMmx1's odour response in substitution experiments. *a* and *a'*, The phasic odour response is reduced after backward pairing in the test compared with pre-test (*a*) but is increased after forward pairing to a long-lasting excitation (*a'*). *b* and *b'*, Post-stimulus interval histogram from 15 consecutive 1-s intervals after stimulus onset. The spike frequency (mean, s.e.m.) in each interval is calculated as the number of spikes in that interval minus the spontaneous fre-

quency 10 s before stimulus onset. Baseline activity is shown as a dashed line at 0 Hz. The excitation of VUMmx1 in the forward paradigm drops to baseline during interval 16 to 20 (large black square in *b'*). The total number of spikes (spike number minus baseline) in 15 s after stimulus onset increases after forward pairing for the test compared with pre-test, but is unchanged following backward-pairing (see insets).

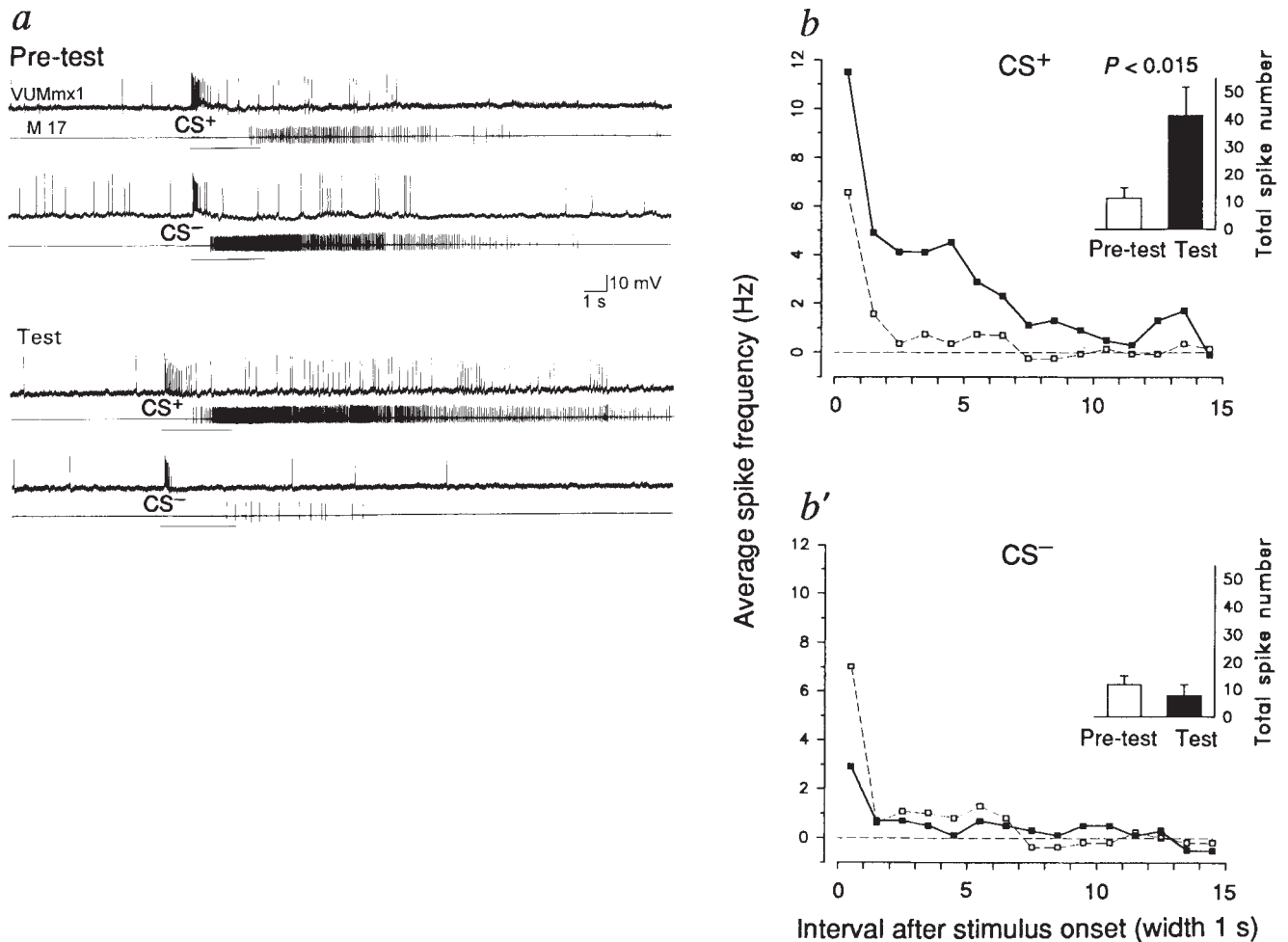


FIG. 4 Response plasticity of VUMmx1 in differential conditioning experiments ($n=5$). Bees were differentially conditioned to two odours (carnation or orange-blossom) in 5 trials. One odour was specifically paired with the US (CS^+ ; intertrial interval, 1 min). The other (CS^-) was delivered specifically unpaired inbetween. Selection of odours as CS^+ or CS^- was exchanged in consecutive experiments. **a**, Responses recorded from VUMmx1 and M17 5 min before (pre-test) and 10 min after (test) differential conditioning. A spontaneous PER to both odours is evoked during pre-test, as indicated by EMGs. Test responses of VUMmx1 and M17 to CS^+ are increased in strength and duration, whereas responses

to CS^- are decreased (M17) or remain unchanged (VUMmx1). **b** and **b'**, Post-stimulus interval histogram reveals that the significant increase in VUMmx1's response to the CS^+ in the test as compared with pre-test (inset) is due to a prolonged excitation (b). The small drop in the total spike number for the CS^- (inset) is due to a reduction of the phasic excitation of VUMmx1 (**b'**). A second test (5–10 min after the test), with reversed order of odours to exclude extinction-like effects between successive odour stimulations, reveals the same results. Spike numbers were calculated as for Fig. 3.

Received 16 April; accepted 13 September 1993.

- Rescorla, R. A. *Rev. Neurosci.* **11**, 329–352 (1988).
- Rescorla, R. A. & Wagner, A. R. in *Classical Conditioning II: Current Research and Theory* (eds Black, A. H. & Prokasy, W. F.) 64–99 (Appleton-Century-Crofts, New York, 1972).
- Wagner, A. R. in *Information Processing in Animals: Memory Mechanisms* (eds Spear, N. E. & Miller, R. R.) 5–47 (Erlbaum, Hillsdale, N. J., 1981).
- Hawkins, R. D. & Kandel, E. R. *Psychol. Rev.* **91**, 375–391 (1984).
- Menzel, R. & Bitterman, M. E. in *Neuroethology and Behavioral Physiology* (eds Huber, F. & Markel, H.) 206–215 (Springer, Berlin, 1988).
- Menzel, R. in *Neurobiology of Comparative Cognition* (eds Kesner, R. P. & Olton, D. S.) 237–292 (Erlbaum, Hillsdale, NJ, 1990).
- Bitterman, M. E., Menzel, R., Fietz, A. & Schäfer, S. *J. comp. Physiol.* **A97**, 107–119 (1983).
- Smith, B. H. *J. exp. Biol.* **161**, 367–382 (1991).
- Arnold, G., Masson, C. & Budharugsa, S. *Cell Tissue Res.* **242**, 593–605 (1985).
- Mobbs, P. *Phil. Trans. R. Soc. B298*, 309–354 (1982).
- Menzel, R., Greggers, U. & Hammer, M. in *Insect Learning: Ecological and Evolutionary Perspectives* (eds Papaj, D. R. & Lewis, A. C.) 79–125 (Chapman & Hall, New York, London, 1993).
- Rehder, V. *J. Insect Physiol.* **33**, 303–311 (1987).

- Mackey, S. L., Kandel, E. R. & Hawkins, R. D. *J. Neurosci.* **9**, 4227–4235 (1989).
- Mauk, M. D., Steinmetz, J. E. & Thompson, R. F. *Proc. natn. Acad. Sci. USA* **83**, 5349–5353 (1986).
- Farley, J. *Behav. Neurosci.* **101**, 28–56 (1987).
- Erber, J., Mazur, T. & Menzel, R. *Physiol. Entomol.* **5**, 343–358 (1980).
- Mauelshagen, J. *J. Neurophysiol.* **69**, 609–625 (1993).
- Heisenberg, M., Borst, A., Wagner, S. & Byers, D. *J. Neurogenet.* **2**, 1–30 (1985).
- Balling, A., Technau, G. M. & Heisenberg, M. *J. Neurogenet.* **4**, 65–73 (1987).
- Nighorn, A., Healy, M. J. & Davis, R. L. *Neuron* **6**, 455–467 (1991).
- Han, P.-L., Levin, L. R., Reed, R. R. & Davis, R. L. *Neuron* **9**, 619–627 (1992).
- Buonomano, D. V., Baxter, D. A. & Byrne, J. H. *Neural Networks* **3**, 507–523 (1990).
- Hawkins, R. D. in *Computational Models of Learning in Simple Neural Systems* (eds Hawkins, R. D. & Bower, G. H.) 65–108 (Academic, San Diego, 1989).
- Rybick, J. & Menzel, R. *J. comp. Neurol.* **334**, 444–465 (1993).
- Rehder, V. *J. comp. Neurol.* **279**, 499–513 (1989).
- Strausfeld, N. *Atlas of an Insect Brain* (Springer, Berlin, Heidelberg, New York, 1976).

ACKNOWLEDGEMENTS. I thank R. Menzel for support and encouragement, and R. Menzel and especially P. Stevenson for their comments on the manuscript.

On the molecular origin of photoreceptor noise

Robert B. Barlow^{*§}, Robert R. Birge[†],
Ehud Kaplan^{‡§} & Jack R. Tallent[†]

^{*} Institute for Sensory Research and Department of Bioengineering and Neuroscience, and [†] Department of Chemistry, Syracuse University, Syracuse, New York 13244, USA

[‡] The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

[§] Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

RETINAL photoreceptors are noisy. They generate discrete electrical events in the dark indistinguishable from those evoked by light^{1,2} and thereby limit visual sensitivity at low levels of illumination^{3,4}. The random spontaneous events are strongly temperature-dependent and have been attributed to thermal isomerizations of the vitamin A chromophore of rhodopsin, the light-sensitive molecule in photoreceptors^{1,5,6}. But thermal generation of dark events in both vertebrate and invertebrate photoreceptors requires activation energies in the range of 23 to 27 kcal mol⁻¹, which are significantly less than the energy barrier of 45 kcal mol⁻¹ for photoisomerization of the chromophore of native rhodopsin⁷⁻⁹. We propose that photoreceptor noise results from the thermal isomerization of a relatively unstable form of rhodopsin, one in which the Schiff-base linkage between the chromophore and protein is unprotonated. This molecular mechanism is supported by both theoretical calculations of the properties of rhodopsin and experimental measurements of the properties of photoreceptor noise.

We can see faint stars and distinguish objects in direct sunlight giving us an operating range of more than 10 log units of light intensity. Our ability to detect dim lights is limited by noise in our visual system. For us and for many animals, retinal photoreceptors appear to be the primary source of noise that limits visual performance. Such 'dark' noise has been attributed to the thermal isomerization of the chromophore of rhodopsin^{5,6}. This explanation, however, is not consistent with the energetics of the ground-state isomerization of the protonated Schiff-base chromophore of native rhodopsin⁷⁻⁹.

The principal mode of activation of rhodopsin is photochemical. When the 11-*cis* isomer of native rhodopsin absorbs green light ($\lambda_{\text{max}} \sim 500$ nm), the molecule is excited to the Frank-Condon region of the first excited, singlet-state surface (R^{FC} in Fig. 1), from which it can isomerize to the all-*trans* chromophore of bathorhodopsin, a barrierless process which accounts for the high quantum efficiency of the primary event (67%)⁷. The ground-state process, however, entails a high energy barrier of ~ 45 kcal mol⁻¹, which is far higher than the Arrhenius activation energy of ~ 27 kcal mol⁻¹ we measure for photoreceptor noise in the *Limulus* lateral eye (Fig. 2). It is also higher than the Arrhenius energies of 21.9 ± 1.6 kcal mol⁻¹ for toad rods¹ and 36 kcal mol⁻¹ for dogfish rods¹⁰. In all cases, the thermal isomerization of the chromophore in native rhodopsin cannot account for photoreceptor dark noise. In addition, chemical, thermodynamic and kinetic analyses¹¹⁻¹³ as well as site-directed mutagenesis^{13,14} indicate that denaturation, spontaneous conformational changes, or deprotonation of the chromophore without isomerization cannot activate rhodopsin in the dark.

We propose that a two-step ground-state process generates photoreceptor dark noise. Step one is the deprotonation of the

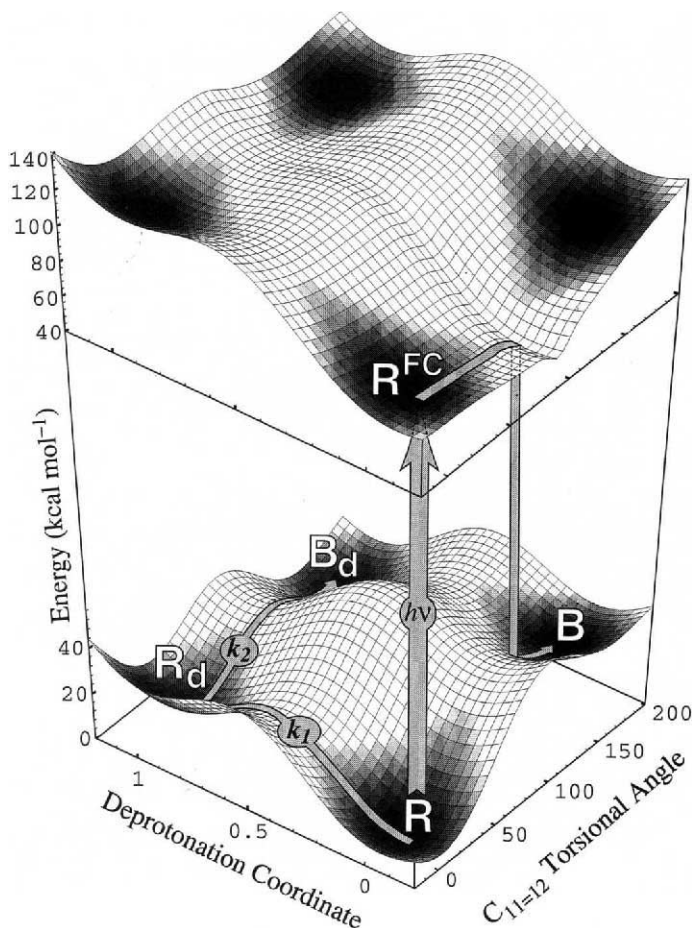


FIG. 1 Photochemical and thermal pathways of activation of rhodopsin. Shown are the ground- and first excited singlet-state adiabatic potential surfaces of the protein-bound chromophore as a function of two degrees of freedom, isomerization of rhodopsin from 11-*cis* to the 11-*trans* form and deprotonation of the Schiff-base linkage. The surfaces are calculated from MNDO/AM1 and INDO-PSDCI molecular orbital theory as described in refs 7 and 8, with the addition of three water molecules to the simulation of the binding site and the examination of the deprotonation coordinate. We normalized this coordinate such that '0' represents the Schiff-base (N-H) bond length at equilibrium and '1' represents the N-H⁺(H₂O)₃ separation at equilibrium solvation of the proton. Spectroscopic and thermochemical studies⁷ support the assumption that water is present in the binding site. The potential surfaces are calculated by minimizing the geometry of the chromophore and three water molecules as a function of two coordinates, the C₁₁-C₁₂ double bond torsional angle (0° = 11-*cis*, 180° = 11-*trans*) and the deprotonation coordinate (0, covalently bonded; 1, deprotonated). The key ground-state minima and calculated relative energies (kcal mol⁻¹) are as follows: R (rhodopsin, $\Delta E = 0$), B (bathorhodopsin, $\Delta E = 31$), R_d (rhodopsin with a deprotonated 11-*cis* chromophore, $\Delta E = 17$), B_d (bathorhodopsin with a deprotonated 11-*trans* chromophore, $\Delta E = 23$). The corresponding excited-state potential surface is shifted vertically to help viewing of the ground-state surface. R^{FC} designates the Franck-Condon excited state of rhodopsin. The photochemical activation of rhodopsin proceeds by excitation into R^{FC}, partial isomerization in the excited state to a C₁₁-C₁₂ distorted geometry ($\sim 90^\circ$), crossing into the ground state and continued rotation about the C₁₁-C₁₂ to form bathorhodopsin (B_d)⁸. The thermal activation of rhodopsin involves a two-step process: deprotonation of the chromophore (rate = $k_1 = A_1 \exp(-27/RT)$) followed by isomerization of deprotonated chromophore (rate = $k_2 = A_2 \exp(-23/RT)$), where simple Arrhenius-rate equations are assumed with prefactors (unknown) of A₁ and A₂ and activation energies in kcal mol⁻¹. Activation to the first excited singlet state of R_d requires absorption of higher energy photons ($\lambda_{\text{max}} = 380$ nm, 73 kcal mol⁻¹) than for native rhodopsin, but the ground-state barrier for isomerization of 23 kcal mol⁻¹ is considerably lower than for native rhodopsin, making the unprotonated form relatively less stable. Energy profiles are based on models of the bovine rhodopsin binding site which has 49% homology (C. Smith, personal communication) with *Limulus* rhodopsin³⁰. Both rhodopsins contain a lysine residue at location 296, a retinal binding site, and glutamate at location 113, the potential Schiff-base counterion.

11-*cis* Schiff-base linkage, and step two is the thermal isomerization of 11-*cis* retinal to the all-*trans* form. Figure 1 shows the calculated energy profiles of the ground and excited states of rhodopsin as a function of isomerization and deprotonation. Deprotonation destabilizes the protein in the ground state by lowering the energy barrier for isomerization of the chromophore from $\sim 45 \text{ kcal mol}^{-1}$ to $\sim 23 \text{ kcal mol}^{-1}$. Reaction rate theory suggests that the overall two-step process (k_1 and k_2 in Fig. 1) would be experimentally indistinguishable from a one-step process with an 'apparent activation energy' of $27\text{--}30 \text{ kcal mol}^{-1}$. Although our theoretical calculations are approximate because of assumptions regarding the geometry of the binding site⁷ and use of semiempirical molecular orbital theory⁸, the calculations provide a perspective on the relative viability of proposed mechanisms for thermal activation of rho-

dopsin. The calculations suggest that thermal isomerization of the protonated Schiff-based chromophore cannot be a dominant mode of activation. In contrast, a two-step process involving deprotonation followed by isomerization yields a theoretical activation energy ($\sim 27 \text{ kcal mol}^{-1}$) that is consistent with activation energies measured for thermal noise in photoreceptors ($22\text{--}36 \text{ kcal mol}^{-1}$).

Further evidence that a two-step process generates photoreceptor noise comes from an analysis of the circadian rhythm

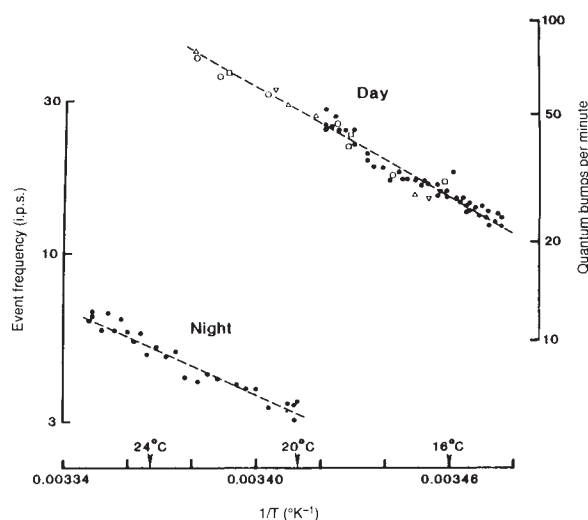


FIG. 2 Temperature influence on the frequency of occurrence of dark events in the *Limulus* retina. Shown are Arrhenius plots of event frequencies on log scales against reciprocal absolute temperature. Filled circles give the spike activity of ~ 25 optic nerve fibres recorded *in vivo* with a tungsten electrode inserted in the lateral optic nerve trunk of a restrained animal immersed in a sea-water aquarium located in a darkened, shielded cage. A thermistor held in contact with the cornea monitored lateral-eye temperature which was shifted in the range of 14° to 26°C by heat-exchange coils located in the aquarium and controlled by an external heater-cooler. During the day, spontaneous optic nerve activity was $28 \text{ impulses s}^{-1}$ (or $\sim 1 \text{ impulses s}^{-1}$ per optic nerve fibre) with eye temperature set at 20°C . A 6°C decrease to 14°C reduced spontaneous spike activity from 28 to 12 impulses s^{-1} (left ordinate). At night the clock's input reduced spontaneous activity to 3 impulses s^{-1} , and increasing eye temperature from 20 to 26°C more than doubled the rate to 6.5 impulses s^{-1} . Eye temperature was changed at a rate of 0.6 to $1.0^\circ \text{C h}^{-1}$ both day and night. Unfilled symbols give the results of four measurements from as many eyes of the temperature effect on the frequency of occurrence of quantum bumps recorded intracellularly in darkness from single photoreceptor cells with glass microelectrodes. We studied the thermal properties of quantum bump noise in single photoreceptor cells *in vitro* because changing the temperature of the eye in the animal elicited muscular movements which disrupted intracellular recordings from single cells. Excising the eye removed the clock's influence and yielded high spontaneous activity characteristic of the daytime state of the eye. At 20°C the spontaneous activity was $\sim 1 \text{ bump s}^{-1}$ per cell. The rates of quantum bump production and daytime optic nerve activity have nearly the same steep temperature dependence. Multiplying the slope of the dashed line for 'Day' data by the gas constant yields an apparent activation energy of $25.5 \text{ kcal mol}^{-1}$. The average activation energy is $26.3 \pm 7.8 \text{ kcal mol}^{-1}$ ($n=10$) for optic nerve impulses and 26.5 ± 7.5 ($n=8$) for quantum bumps. The dashed line for 'Night' data yields an apparent activation energy of $21.2 \text{ kcal mol}^{-1}$ with an average value of $27.9 \pm 6.5 \text{ kcal mol}^{-1}$ ($n=12$).

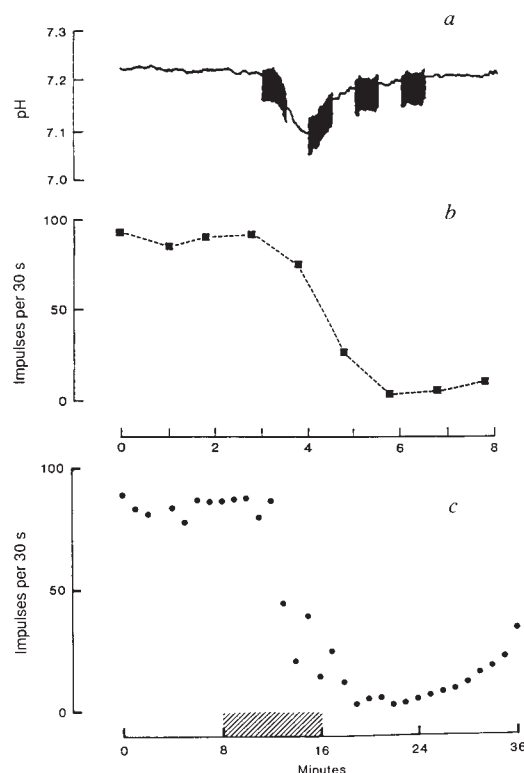


FIG. 3 Efferent optic nerve activity reduces both pH and dark noise of the *Limulus* retina. *a*, Voltage trace shows retinal pH recorded *in vivo* with an H^+ -selective microelectrode (tip diameter $\sim 4 \mu\text{m}$; ref. 31) inserted into retinal tissue by an access hole cut in the cornea. The technique follows that developed for intracellular recordings of retinal cells *in vivo*¹⁸ except that the tip of the pH electrode was located in extracellular retinal space. The optic nerve trunk was exposed, cut, and its distal end was pulled into a suction electrode. The nerve trunk was stimulated beginning at 3 min with a 30-s train of current pulses (2 s^{-1}) repeated three times to mimic the clock's efferent input to the eye³². The current pulses produced artefacts (dark bands) on the microelectrode recording. The first volley of current pulses decreased retinal pH from 7.23 to 7.09, but subsequent volleys over the next 3 min had no effect. The pH changes could be replicated only after the preparation remained unstimulated for $> 45 \text{ min}$. Decreases in retinal pH were detected in 6 out of 22 experiments. This experiment yielded the largest change of 0.14 pH units, and the average for all 6 experiments was 0.06 pH units. The time course of pH changes for each experiment was similar to that shown here. Not all experiments exhibited the same pH changes perhaps because of retinal buffering and variation of location of the tip of the pH microelectrode in retinal tissue. *b*, Effect of efferent input on the spontaneous impulse activity of three lateral optic nerve fibres of an animal maintained in darkness. Condition of the animal and temporal sequence of efferent activity match those of the pH experiment in *a*. Ordinate gives the number of impulses recorded in 30-s bins every min. *c*, Effect of low retinal pH on the spontaneous activity of three optic nerve fibres *in situ*. This experiment was done on a different animal from that in *b* but the conditions were identical. Beginning at 8 min, $16 \mu\text{l}$ of saline (pH 6.86) was injected at the rate of $2 \mu\text{l min}^{-1}$ for a total of 8 min. Spontaneous activity decreased 5 min after injection began ($t=13 \text{ min}$), reached a minimum at $t=22 \text{ min}$, and returned to the preinjection level at $\sim 45 \text{ min}$.

of dark noise in the *Limulus* retina. At night, efferent nerve signals from a circadian clock in the brain reduce the rate of spontaneous events from photoreceptors and second-order cells without affecting those evoked by light^{15,16}. The Arrhenius energies for noise generation are nearly equal both day and night (Fig. 2), indicating that the clock does not reduce noise by increasing the energy required to generate it.

We propose that dark noise in *Limulus* photoreceptors results from the thermal isomerization of a small (<0.01%) population¹⁷ of rhodopsin molecules containing unprotonated Schiff-base chromophores. At night the clock's efferent input reduces noise by decreasing the number of molecules in the unprotonated state and does so by lowering external pH in the vicinity of the rhodopsin-containing membrane.

Figure 3 presents tests of this molecular mechanism. Mimicking the clock's efferent input by delivering current shocks to the optic nerve decreases retinal pH by 0.14 units (Fig. 3a). Mimicking the efferent input as in Fig. 3a reduces retinal dark noise as measured by spontaneous optic nerve activity (Fig. 3b). The reduction of noise follows by ~1 min the transient decrease in pH. Why the noise changes outlast the pH changes is not clear. Injecting the eye with mildly acidic saline (pH 6.86) reduces photoreceptor dark noise ~50 times, which corresponds well with the observations in *Limulus* ventral photoreceptors of noise reduction by low pH¹⁸ and the accessibility of visual pigment to changes in extracellular pH¹⁹ (Fig. 3c). Neural activity lowers pH in other tissues²⁰. Efferent inputs to the *Limulus* eye terminate directly on rhodopsin-containing membranes of each photoreceptor²¹, but how they can lower pH is not known. The source may be presynaptic activity because agonists of the efferent input that act postsynaptically do not strongly influence photoreceptor noise^{22,23}. In summary, the results of these physiological tests are consistent with the molecular mechanism we propose.

Horseshoe crabs have evolved circadian mechanisms for increasing the sensitivity of their lateral eyes at night, apparently to help them find mates at the water's edge^{24,25}. With a relatively high content of rhodopsin²¹, *Limulus* photoreceptors are highly sensitive but noisy. The theoretical and experimental studies we report here support the idea that the circadian efferent input to the photoreceptors reduces their noise by reducing the small population of relatively unstable, unprotonated rhodopsin molecules.

Whether other animals modulate photoreceptor dark noise as we propose for *Limulus* is not known. The near proportionality between dark noise and rhodopsin content of photoreceptors in a variety of animals^{1,10,26,27} points to rhodopsin as the noise source. Some animals such as insects²⁷ and primates²⁶ may minimize dark noise by possessing relatively small photoreceptors with correspondingly small amounts of rhodopsin. Animals such as toads and other amphibians with large photoreceptors and large amounts of rhodopsin per cell may minimize dark noise by living at relatively low temperatures⁴. Finally, the pH of some vertebrate photoreceptors decreases under prolonged dark adaptation^{28,29}, leaving open the possibility that rhodopsin stabilization contributes to high visual sensitivity in these animals as it appears to in *Limulus*. □

Received 15 June; accepted 24 August 1993.

1. Baylor, D. A., Matthews, G. & Yau, K. W. *J. Physiol.* **309**, 591–621 (1980).
2. Yeandle, S. Am. *J. Ophthalmol.* **46**, 82–87 (1958).
3. Copenhagen, D. R., Donner, K. & Reuter, T. *J. J. Physiol.* **393**, 667–680 (1987).
4. Aho, A. C., Donner, K., Hyden, C., Larsen, L. O. & Reuter, T. *Nature* **334**, 348–350 (1988).
5. Barlow, H. B. *J. Opt. Soc. Am.* **46**, 634–639 (1956).
6. Barlow, H. B. *Nature* **334**, 296–297 (1988).
7. Birge, R. R. *Biochim. biophys. Acta* **1016**, 293–327 (1990).
8. Tallent, J. R., Hyde, E. W., Findsen, L. A., Fox, G. C. & Birge, R. R. *J. Am. chem. Soc.* **114**, 1581–1592 (1992).
9. Schick, G. A., Cooper, T. M., Holloway, R. A., Murray, L. P. & Birge, R. R. *Biochemistry* **26**, 2556–2562 (1987).
10. Ashmore, J. F. & Falk, G. *J. Physiol.* **332**, 272–297 (1977).
11. Hubbard, R. *J. Biol. Chem.* **241**, 1814–1818 (1966).
12. Longstaff, C., Calhoun, R. D. & Rando, R. R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4209–4213 (1986).

13. Sakmar, T. P., Franke, R. R. & Khorana, H. G. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8309–8313 (1989).
14. Fahmy, K. & Sakmar, T. P. *Biochemistry* **32**, 9165–9171 (1993).
15. Barlow, R. B. Jr, Kaplan, E., Renninger, G. H. & Saito, T. *J. gen. Physiol.* **89**, 353–378 (1987).
16. Kaplan, E. & Barlow, R. B. Jr *Nature* **286**, 393–395 (1980).
17. Birge, R. R. *Biophys. J.* **64**, 1371–1372 (1993).
18. Corson, D. W. & Fein, A. *Brain Res.* **193**, 558–561 (1980).
19. Lisman, J. E. & Strong, L. A. *J. gen. Physiol.* **73**, 219–243 (1980).
20. Chesler, M. *Prog. Neurobiol.* **34**, 401–427 (1990).
21. Fahrenbach, W. H. Z. *zellforsch. mikrosk. anat.* **93**, 451–483 (1969).
22. Schneider, M. H., Lehman, H. K. & Barlow, R. B. Jr, *Invest. Ophth. Vis. Sci.* **28**, (suppl.) 186 (1987).
23. Kass, L., Pelletier, J. L., Renninger, G. H. & Barlow, R. B. Jr *J. comp. Physiol.* **A164**, 95–105 (1988).
24. Barlow, R. B. Jr, Ireland, L. C. & Kass, L. *Nature* **296**, 65–66 (1982).
25. Powers, M. K., Barlow, R. B. Jr & Kass, L. *Visual Neurosci.* **7**, 179–189 (1991).
26. Baylor, D. A., Nunn, B. J. & Schnapf, J. L. *J. Physiol.* **357**, 575–607 (1984).
27. Lillywhite, P. G. *J. comp. Physiol.* **122**, 189–200 (1977).
28. Borja, G. A., Karwowski, C. J. & Steinberg, R. H. *Vision Res.* **29**, 1069–1077 (1989).
29. Oakley, B. II & Wen, R. J. *Physiol.* **419**, 353–378 (1989).
30. Smith, W. C., Greenberg, R. M., Price, D. A. & Battelle, B. A. *Invest. Ophth. Vis. Sci.* **33**, (suppl.) 1004 (1992).
31. Chesler, M. & Kraig, R. P. *Am. J. Physiol.* **253**, R666–R670 (1987).
32. Barlow, R. B. Jr *J. Neurosci.* **3**, 856–870 (1983).

ACKNOWLEDGEMENTS. We thank D. A. Baylor, P. A. Hargrave and T. P. Sakmar for discussions, and M. Chesler and T. H. Silbaugh for assistance. This research was supported by the National Science Foundation, the National Institutes of General Medical Science and the National Eye Institute.

Fast axonal transport is required for growth cone advance

Chris Martenson, Kimberly Stone, Mary Reedy & Michael Sheetz*

Departments of Pathology and Cell Biology, Duke University Medical Center, Durham, North Carolina 27710, USA

GROWTH CONES are capable of advancing despite linkage to a stationary axonal cytoskeleton in chick and murine dorsal root ganglion neurites^{1,2}. Several lines of evidence point to the growth cone as the site of cytoskeletal elongation^{3,4}. Fast axonal transport is probably the means by which cytoskeletal elements⁵ or cofactors are rapidly moved through the axon. We report that direct, but reversible, inhibition of fast axonal transport with laser optical tweezers inhibits growth cone motility if cytoskeletal attachment to the cell body is maintained. Advancement ceases after a distance-dependent lag period which correlates with the rate of fast axonal transport. But severing the axonal cytoskeleton with the laser tweezers allows growth cones to advance considerably further. We suggest that axon elongation requires fast axonal transport but growth cone motility does not.

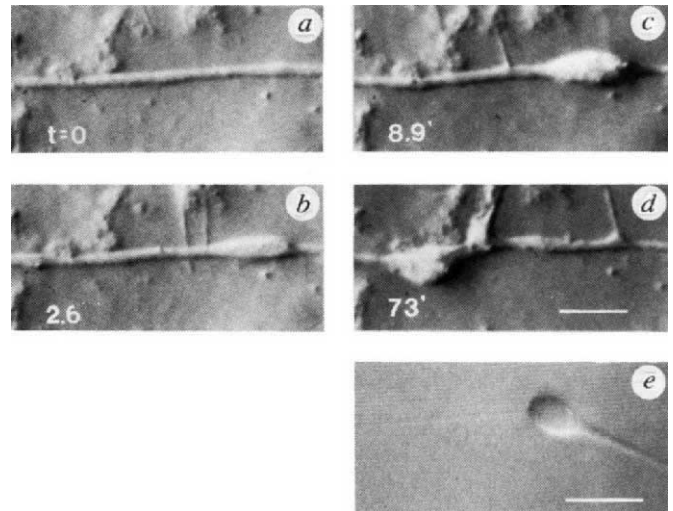
In previous studies, when fast axonal transport was ligated by mechanical axonal transection, growth cones collapsed but then re-extended before halting^{6,7}. Using optical tweezers capable of inhibiting the movement of intracellular vesicles non-invasively⁸, we determined how growth cones respond when deprived of fast transport vesicles while remaining cytoskeletally linked to the cell body. Failure to advance under these conditions would suggest that factors carried by the fast transport system to the vicinity of the growth cone mediate the process of cytoskeletal elongation.

Directing the optical tweezers at a large vesicle, such as a mitochondria or a 1–2 µm intra-axonal varicosity, causes complete occlusion to vesicular transport in 1–2 min (Fig. 1). Initially, only large vesicles are halted by the tweezers but, as they accumulate, progressively smaller vesicles become trapped by the block. The result is an axonal swelling at the focal point of the optical tweezers. The rate at which the swellings induced by optical tweezers increased in diameter, and the final diameter

* To whom correspondence should be addressed.

FIG. 1 Axonal blockage induced by laser tweezers, and subsequent recovery. *a*, Axon is 1.1 μm in diameter. Laser irradiation was for 5.6 min at 200 mW. No observable vesicular traffic traversed the block after 1.8 min of laser irradiation. Anterograde traffic proceeds from left to right. *b*, Same axon after 2.6 min. Growth of the bulge correlates with the influx volume of vesicles. *c*, Axon after 8.9 min. Vesicles accumulate for another 39 s, at which point they approach and traverse the blockage. *d*, Same axon after 73 min. The swelling has decreased in diameter and moved 18 μm retrogradely. Scale bar, 5 μm . *e*, View of laser-transected axon 5 min after 1.2 min of 200-mW laser irradiation. The diameter of axon seen to the right of the axoplasm 'bead' is comparable to the original diameter of the axon. Vesicle movements can be observed to the right of the 'bead', indicating membrane integrity. This bead of axoplasm withdrew towards the growth cone which is 124 μm to the right. The membranous tether is estimated to be 40–80 nm in diameter. Scale bar, 5 μm .

METHODS. Dorsal root ganglion explants were dissected from 12-day chick embryos and plated on poly-D-lysine-coated size 0 coverslips which had been exposed for 15 min to a 1:50 Matrigel:MEME solution. MEME without phenol red (GIBCO) was used as a base stock with 600 mg % (w/v) glucose, 20 mM HEPES buffer, 2 mM glutamine, 50 units per ml of penicillin and streptomycin, N2 supplement (GIBCO), and 75 mg ml⁻¹ NGF. All experiments were illuminated by a mercury-arc halogen lamp passed through an ultraviolet, an infrared, and a 610-nm high-pass filter element. Three growth cones were tracked for 45 min each under this illumination and showed no observable defects in either morphology or rate of advance. Details of this laser-tweezers microscope system have been described¹⁴.



of the swellings, correlated with the volume of axonal vesicular traffic. Electron micrographs of swellings induced by laser tweezers reveal an abundance of membranous profiles and continuous cytoskeletal elements throughout the block (Fig. 2). Vesicles neither stopped nor accumulated after defocusing the optical tweezers into a poor trapping configuration, indicating that the swellings are generated by the interaction between optical tweezers and vesicle and not by the dose of photons delivered.

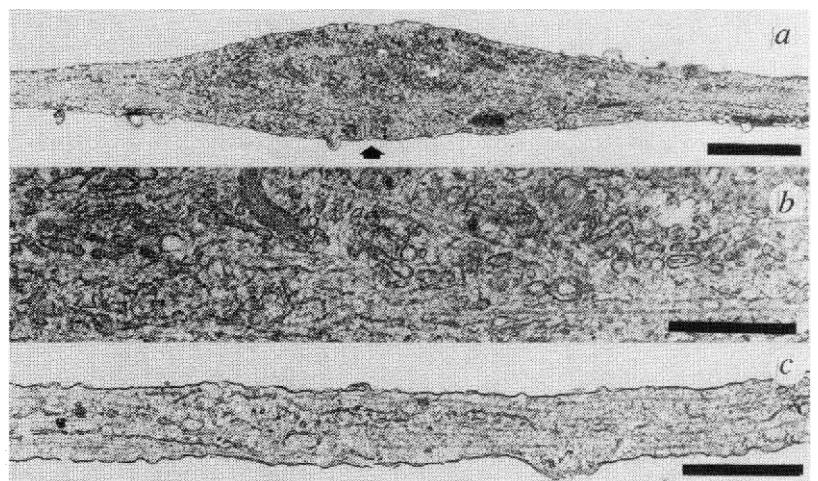
Blocking vesicular transport within an axon stopped growth cone advance in all 29 growth cones investigated. Growth cones

>190 μm from the block continued to advance during the blocking procedure and stopped abruptly after a lag period (Fig. 3a). The length of this lag period was linearly dependent on the distance between the block and the growth cone, with an average signalling rate of 1.28 $\mu\text{m s}^{-1}$ (Fig. 3d). For ~70% of growth cones cessation of advance correlated with morphological alterations, characterized by a collapse of the lamellar regions and an increase in the length and number of filopodia.

The tweezers-induced blocks are not the result of irreversible damage because resumption of vesicular traffic across the block

FIG. 2 Electron micrograph of laser-blocked axon. *a*, Axon was blocked for 1.5 min at 200 mW, with the centre of the beam focused at the location indicated by the arrowhead. Cell body is located to the left. The preponderance of trapped membrane is represented by endoplasmic reticulum-like tubulo-vesicular profiles. Similar profiles are seen in the axon on both sides of the block and in control axon sections. Integrity of axonal membrane, internal membranous features, and abundance of mitochondrial profiles indicate the axonal membrane remained intact throughout the procedure. Two other laser blocks, fixed with glutaraldehyde/tannic acid/osmium and embedded in araldite, also revealed accumulation of similar organelles. Scale bar, 2 μm . *b*, Higher magnification view of centre of same block. As in endogenous swellings on non-blocked axons, the microtubules and neurofilaments that traverse the block region are often distorted from a parallel orientation, probably being forced aside by the densely packed vesicles. Scale bar, 1 μm . *c*, Section of control axon reveals comparable membranous organelle profiles and preservation of axonal membrane. Scale bar, 1 μm .

METHODS. All laser irradiation and cell culture protocols are described in legend to Fig. 1, except that coverslips were first carbon-coated over electron microscope finder grids and ultraviolet light-sterilized. Immediately following block formation, the sample on the heated stage was fixed by flowing ~300–400 μl of 3% glutaraldehyde freshly diluted into the growth medium, pH 7.4. Post-fixation was in 1% OsO₄ in 0.2 M



collidine buffer at 4 °C for 30 min. Sample was block-stained in 2% aqueous uranyl acetate for 30 min, dehydrated, and embedded in Epon. 70-nm longitudinal sections were collected in series and post-stained with uranyl acetate and lead citrate. Micrographs were photographed on a Phillips CM-10 transmission microscope.

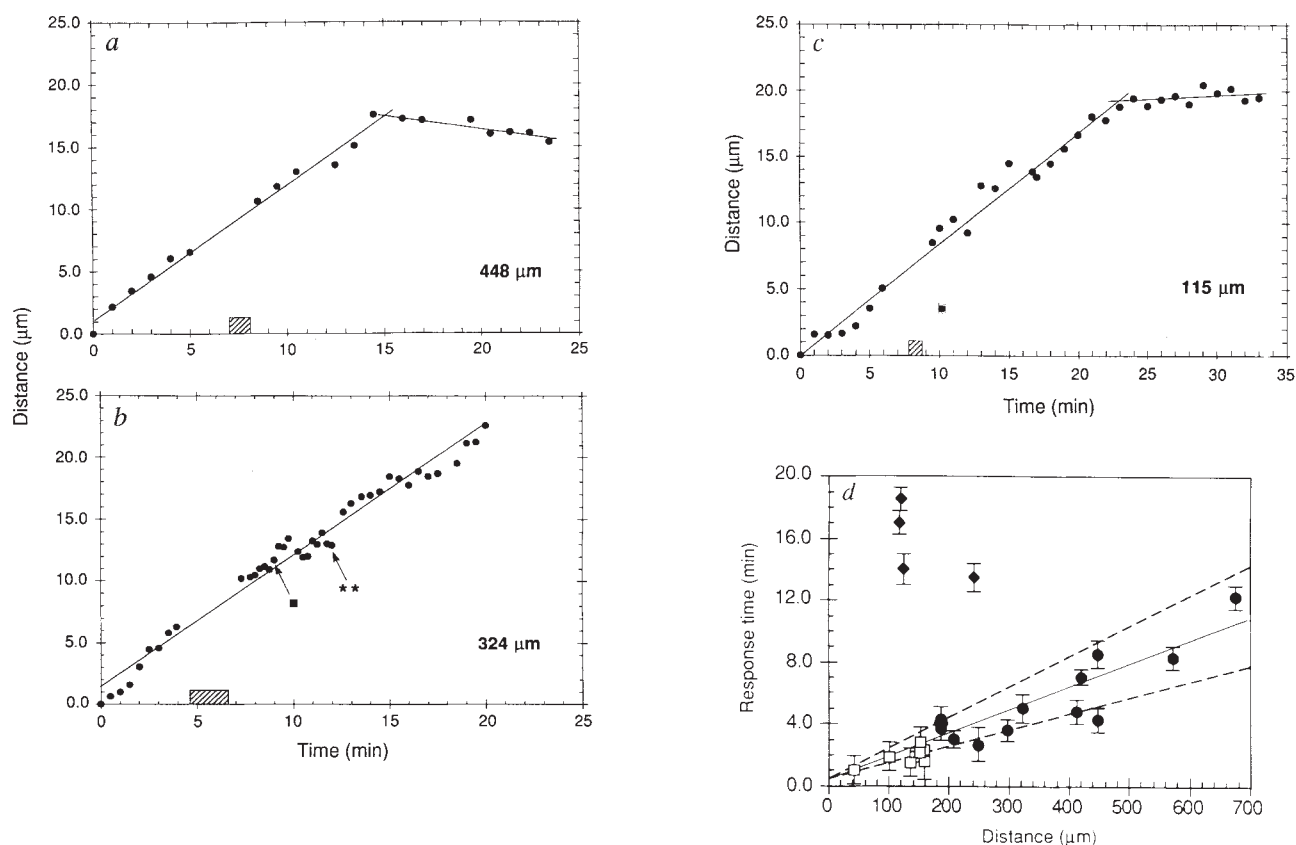


FIG. 3 *a*, Growth cone centroids measured before and after axon blockage. Laser irradiation 448 μm from the growth cone lasted 61 s at 200 mW (hatched rectangle on the ordinate). Growth cone advances for 8.5 min after the start of the optical-tweezing procedure indicating a signalling rate of $0.88 \mu\text{m s}^{-1}$. Collapse of lamellar structures and cessation of growth cone advance occur at the point where the two slopes intersect at which a slow retraction commences. *b*, Growth cone centroids measured before and after blockage of the axon 324 μm away. Laser irradiation was for 1.9 min at 200 mW (hatched rectangle) and ceased the moment blockage was complete. The growth cone advances for 5.5 min from the start of lasering, indicating a signalling rate of $0.98 \mu\text{m s}^{-1}$, loses most of its lamellar area (solid square), and ceases to advance for 2.25 min. Double asterisk indicates the resumption of advance, which in this case coincides with the sudden return of active lamellae. The forward weight provided by the lamellae accounts for the apparent forward jump of the growth cone. The post-recovery advance rate of $0.8 \mu\text{m min}^{-1}$, although less than that of the pre-laser irradiation period, is comparable to the measured rate of most other control growth cones. *c*, Growth cone centroids after laser transection of the axon. Hatched rectangle on the ordinate indicates the period of laser irradiation at 200 mW and 115 μm from the growth cone. A rapid bidirectional retreat of axonal contents at the block site occurred after 1.0 min of irradiation. Growth cone continues to advance for 16.2 min after the start of laser irradiation. A non-transected growth cone would

have ceased moving in 3 min. *d*, Growth-cone advance after laser irradiation for non-severed (filled circles), and transected (filled triangles) growth cones. Open squares represent growth cones that did not significantly advance. The slope of the best-fit line (solid line) through the non-severed data points indicates a signalling rate of $1.28 \mu\text{m s}^{-1}$, with a 99% confidence interval of the slope (dotted lines), bounded by rates of $1.49 \mu\text{m s}^{-1}$ and $0.82 \mu\text{m s}^{-1}$. A diffusion-driven response would operate as the square of the distance. Thus, a generated toxin diffusing through medium at $1 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ would require at least 15 min to travel 300 μm and at least 81 min to travel 700 μm . Additionally, a diffusing toxin would be diluted by the cube of the distance. A toxin diffusing in the plane of a membrane at $1 \times 10^{-8} \text{ cm}^2 \text{ s}^{-1}$ would require 10 times longer to travel equivalent distances. Error bars for all filled data points indicate the length of the lasering procedure. Error bars for the open squares indicate the length of time from the mid-point of the lasering procedure to the re-imaging of the growth cone.

METHODS. All experiments were recorded on a VHS videotape. Stored images were then digitized by a Scion 8-bit video board and stored as PICT files by IMAGE 1.47 (NIH) software. Each growth cone image was outlined and the x and y centroids calculated by Signal Analytic's IPLABS program. All structures of the growth cone were included in the outline. The neck of the growth-cone was arbitrarily set to exist at the point where the axon swelled to 1.5 times its original diameter. No weighting was given to regions of varying optical (or biological) density.

sites occurred routinely after 1–3 min of laser irradiation. Additionally, recovery of normal growth cone motility and morphology was possible if the optical tweezers was turned off when a complete block had formed (Fig. 3*b*). To determine the effects of radiation damage, the optical tweezers were used intermittently (for a cumulative 3 min of laser irradiation) so that no vesicles became trapped and no axonal blockage formed. In these cases, no alterations in growth cone morphology or motility were noted. In addition, focusing the laser $\leq 2 \mu\text{m}$ to the side of an axon for 3 min had no effect, ruling out local heating as a factor. However, recovery of normal growth cone morphology and motility did not correlate directly with the resumption of vesicle transport across the block site, indicating that some form of

damage had occurred. Using longer exposures to laser light (3–5 min), a form of damage to membrane integrity was occasionally observed. In these cases, the growth cones collapsed and withdrew faster than they could be relocated and brought back in focus.

Cytoskeletal breaks, visualized as a rapid bidirectional withdrawal of axoplasm from the block region, sometimes occurred if laser irradiation was conducted within 250 μm of the growth cone. The severed axonal stumps usually remained 'linked' by an extremely thin membranous tether which was fluid and showed no vesicle movements, overt stiffness, or other indications of possessing a cytoskeletal structure (Fig. 1*e*). Severed growth cones continued to advance for at least 10–15 min longer

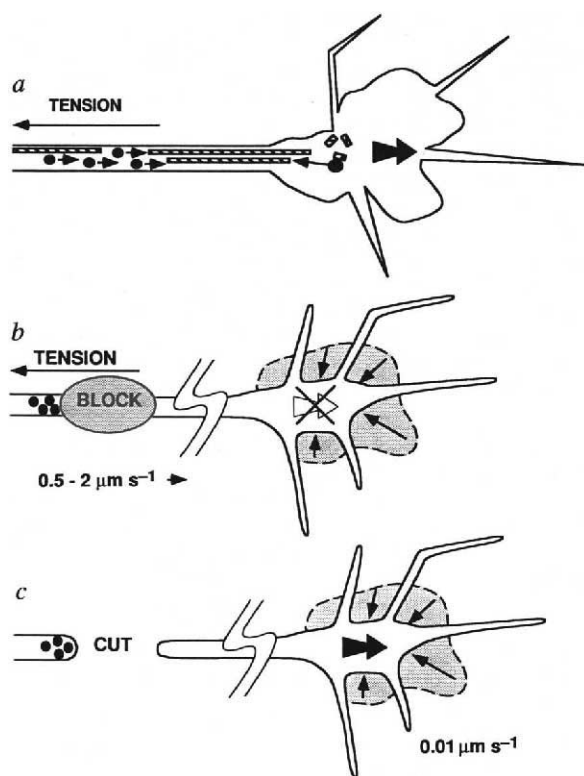


FIG. 4 Summary of growth cone behaviour. *a*, Factors carried by the fast transport system (filled circles) arrive at the growth cone and facilitate the elongation of the cytoskeletal network (hatched rectangles), allowing advance (solid arrow) to occur. *b*, When growth cones are deprived of fast transport material but remain attached to the axonal cytoskeleton, advance is no longer possible (crossed-out arrow) and growth cones collapse. These changes are propagated to the growth cone at the indicated rate of $0.5\text{--}2\ \mu\text{m s}^{-1}$. *c*, When growth cones are transected from the axonal cytoskeletal network and deprived of fast transport material, they collapse but are capable of advancing at the rate indicated. Thus, the only limitation to the forward movement of blocked growth cones is anchorage to the cytoskeleton.

than similarly irradiated non-severed growth cones (Fig. 3c). As both blocked and transected growth cones are deprived of fast transport vesicles to the same extent and have received equivalent doses of laser irradiation, cytoskeletal linkage to the cell body may account for the variance in motile behaviour. The possibility of a slow wave of Ca^{2+} influx propagating down the axon was ruled out as such waves in other systems are much more rapid⁹. Similarly, changes propagated as a result of a blockage of the slow transport system might be expected to travel at the slow transport rate of $0.01\text{--}0.05\ \mu\text{m s}^{-1}$.

The simplest explanation for our findings is that the axonal cytoskeleton provides a mechanical anchor to forward migration (Fig. 4b). For the axonal cytoskeleton to elongate in the absence of a bulk sliding process^{1,2}, cytoskeletal assembly must occur near the growth cone. Fast axonal transport of cytoskeletal elements may occur⁵ but it is also likely that cytoskeletal proteins are conveyed by slow axonal transport (reviewed in ref. 10) and assembly is locally regulated at the growth cone. Such assembly could require fast transport components that either coassemble with the cytoskeleton or are short-lived cofactors in the assembly process (Fig. 4a). Transection (Fig. 4c) would relieve the tension axons are under^{11–13}, allowing the growth cone to advance. □

Received 11 May; accepted 7 September 1993.

- Okabe, S. & Hirokawa, N. *J. Cell Biol.* **107**, 651–664 (1988).
- Okabe, S. & Hirokawa, N. *Nature* **343**, 479–482 (1990).
- Bamburg, J., Bray, D. & Chapman, K. *Nature* **321**, 788–790 (1986).
- Brown, A., Slaughter, T. & Black, M. *J. Cell Biol.* **4**, 867–882 (1992).
- Hollenbeck, P. & Bray, D. *J. Cell Biol.* **105**, 2827–2835 (1987).
- Shaw, G. & Bray, D. *Exp. Cell Res.* **104**, 55–62 (1977).
- Bray, D. *J. Cell Sci.* **37**, 391–410 (1979).
- Ashkin, A., Schutze, K., Dziedzic, J., Euteneuer, U. & Schliwa, M. *Nature* **348**, 346–348 (1990).
- Cornell-Bell, A., Finkbeiner, S., Cooper, M. & Smith, S. *Science* **247**, 470–473 (1990).
- Mitchison, T. & Kirschner, M. *Neuron* **1**, 761–772 (1988).
- Dennerli, T., Joshi, H., Steel, V., Buxbaum, R. E. & Heidemann, S. R. *J. Cell Biol.* **107**, 665–674 (1988).
- Dennerli, T. J., Lamoureux, P., Buxbaum, R. & Heidemann, S. R. *J. Cell Biol.* **109**, 3073–3083 (1989).
- Lamoureux, P., Buxbaum, R. E. & Heidemann, S. R. *Nature* **340**, 159–162 (1989).
- Kucik, D., Kuo, S., Elson, E. & Sheetz, M. *J. Cell Biol.* **114**, 1029–1036 (1991).

ACKNOWLEDGMENTS. We thank S. C. Kuo for technical assistance and for discussion.

Spontaneous loss of T-cell tolerance to glutamic acid decarboxylase in murine insulin-dependent diabetes

Daniel L. Kaufman^{*†‡}, Michael Clare-Salzler^{§||}, Jide Tian^{*}, Thomas Forsthuber^{¶||}, Grace S. P. Ting^{*}, Paul Robinson[§], Mark A. Atkinson[#], Eli E. Sercarz[¶], Allan J. Tobin^{†‡**} & Paul V. Lehmann^{¶||}

^{*} Department of Psychiatry and Biobehavioral Sciences,

[§] Department of Medicine, [¶] Department of Microbiology and Molecular Genetics, ^{**} Department of Biology, [†] Brain Research Institute, and [‡] Molecular Biology Institute, University of California at Los Angeles, Los Angeles, California 90024, USA

[#] Department of Pathology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, Florida 32610, USA

INSULIN-DEPENDENT diabetes mellitus (IDDM) in non-obese diabetic (NOD) mice results from the T-lymphocyte-mediated destruction of the insulin-producing pancreatic β -cells and serves as a model for human IDDM¹. Whereas a number of autoantibodies are associated with IDDM², it is unclear when and to what β -cell antigens pathogenic T cells become activated during the disease process. We report here that a T-helper-1 (Th1) response to glutamate decarboxylase develops in NOD mice at the same time as the onset of insulinitis. This response is initially limited to a confined region of glutamate decarboxylase, but later spreads intramolecularly to additional determinants. Subsequently, T-cell reactivity arises to other β -cell antigens, consistent with intermolecular diversification of the response. Prevention of the spontaneous anti-glutamate decarboxylase response, by tolerization of glutamate decarboxylase-reactive T cells, blocks the development of T-cell autoimmunity to other β -cell antigens, as well as insulinitis and diabetes. Our data suggest that (1) glutamate decarboxylase is a key target antigen in the induction of murine IDDM; (2) autoimmunity to glutamate decarboxylase triggers T-cell responses to other β -cell antigens, and (3) spontaneous autoimmune disease can be prevented by tolerization to the initiating target antigen.

We tested NOD mice from birth to 28 weeks of age for T-cell reactivity to β -cell antigens that are targets of IDDM-associated

|| Present addresses: Department of Pathology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, Florida 32610, USA (M.C.-S.); and Department of Pathology, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA (T.F. and P.V.L.).

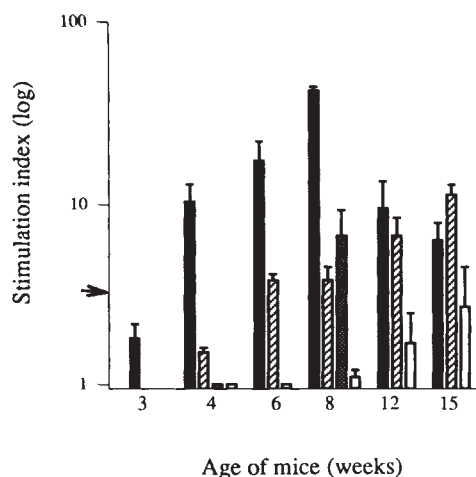


FIG. 1 Proliferative T-cell responses to β -cell antigens develop spontaneously in NOD mice in a defined chronological order. Antigen-induced blastogenesis was measured in spleen cells from newborn to 28-week-old female NOD mice (data from 3–15 weeks are shown). β -Cell antigens include GAD65 (black bars), Hsp65 peptide PALDSLTPANED¹² (striped bars), carboxypeptidase H (grey bar) and insulin (white bars). Data are expressed as stimulation indices (SI) $\pm$ standard error of the mean (s.e.m.), calculated from 3–5 mice tested individually in 2 separate experiments for each time point. Arrow indicates SI = 3, the level of significance. Carboxypeptidase H responses were only tested at 4 and 8 weeks. None of the control antigens (hen egg-white lysozyme, human serum albumin, *E. coli* β -galactosidase or murine myelin basic protein) induced T-cell proliferation at any age. Also, none of the β -cell antigens or control antigens induced proliferation of T cells from age-matched control BALB/c or (NOD $\times$ BALB/c) F₁ mice (data not shown). METHODS. NOD (Taconic farms) and BALB/c mice (Jackson Laboratories) were housed under specific pathogen-free conditions. Spleen cells were tested directly ex vivo for proliferative recall responses. Cells were plated at 1×10^6 cells per well in 96-well microtitre plates in 200 μ l HL-1 medium (Ventrex) containing 2 mM glutamine and 10 μ g ml⁻¹ antigen, or 7 μ M peptide (the optimal concentrations for all time points tested) in triplicate. During the last 16 h of the 72 h culture period, 1 μ Ci [³H]thymidine was added per well. Incorporation of label was measured by liquid scintillation counting. Human GAD65 (ref. 25) and *E. coli* β -galactosidase were both purified from recombinant bacteria using a hexahistidine tag and metal-affinity chromatography²⁶. Bovine carboxypeptidase H was a generous gift from L. Fricker; human insulin was purchased from Eli Lilly.

autoantibodies. These included one of the two forms of glutamate decarboxylase^{3–5} (GAD65, an early target of autoantibodies^{6–9}), carboxypeptidase H¹⁰, insulin² and the immunodominant determinant of heat-shock protein (Hsp65) (refs 11, 12).

Proliferative T-cell responses to these antigens developed spontaneously in a sequential order. First, a response to GAD arose at 4 weeks of age (Fig. 1), concurrent with the onset of insulinitis. This reactivity increased during the next four weeks and then declined to background levels by week 16. At 6 weeks of age, responses to heat-shock protein appeared, increased until week 15 and then diminished (Fig. 1; also ref. 12). Similarly, whereas no response was detected to carboxypeptidase H at 4 weeks old, there was a strong anti-carboxypeptidase H response by week 8. In some mice, a weak response to insulin became detectable at 12–15 weeks old.

The spontaneous development of a proliferative T-cell response to GAD is consistent with, but does not prove, endogenous priming. We therefore tested GAD-reactive T cells for additional properties that distinguish activated/memory lymphocytes from resting/naïve lymphocytes. First, we found that GAD-challenged freshly isolated T cells from 6–9-week-old NOD mice produce interferon- γ (IFN- γ), which is secreted only by preactivated Th1 lymphocytes¹³ (Fig. 2a). Second, whereas the frequency of T cells reactive to control antigens constituted ~ 1 in 10^5 cells in the spleen of 6–9-week-old NOD mice, the frequency of GAD-reactive T cells was almost two orders of magnitude higher, which is consistent with clonal expansion (Fig. 2a). Third, GAD reactivity resided within the L-selectin⁺ fraction of CD4⁺ cells, a phenotype characteristic of activated lymphocytes¹⁴ (Fig. 2b). These findings provide independent evidence that a potentially pathogenic¹⁵ Th1-type T-cell population is spontaneously primed to GAD65 early in the development of NOD diabetes.

We mapped the fine specificity of the anti-GAD T-cell response using an overlapping set of peptides that span GAD65. The initial response, at 4 weeks, involved two adjacent peptides (amino acids 509–528 and 524–543, peptides 34 and 35, respectively; Fig. 3a). During the next 3 weeks, T-cell autoimmunity spread to several additional GAD determinants (including amino acids 246–266, peptide 17, which contains a region of sequence similarity with Coxsackie virus⁹; Fig. 3b, c). Subsequently, reactivity to GAD peptides declined (data not shown), paralleling the loss of response to the whole protein (Fig. 1).

The gradual diversification of the primed autoreactive T-cell repertoire in this naturally occurring autoimmune disease paral-

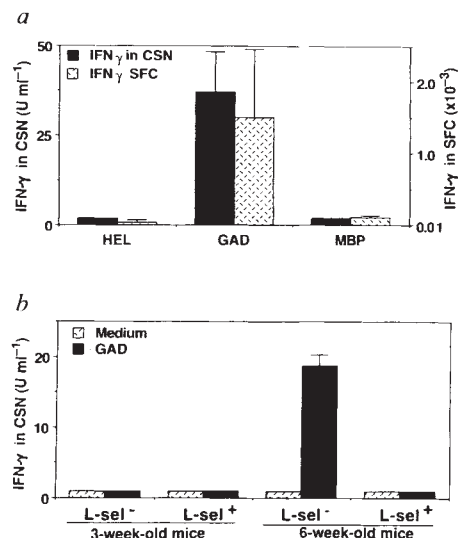


FIG. 2 GAD-specific T cells in 6–9-week-old female NOD mice are primed Th1 type CD4⁺ lymphocytes based on their production of IFN- γ (a and b), enhanced clonal size (a) and L-selectin⁺ phenotype (b). a, Detection of IFN- γ by ELISA in culture supernatants (CSN) of spleen cells from 6–9-week-old mice 48 h after challenge with GAD or control antigens hen egg-white lysozyme (HEL) and murine myelin basic protein (MBP). High concentrations of IFN- γ were detected only in cultures containing GAD. IFN- γ production was measured by ELISA²⁷ using IFN- γ -specific monoclonal antibodies R4-6A2 and XMG 1.2 (Pharmingen). T cells from age-matched BALB/c mice did not respond to GAD or to control antigens (data not shown). The frequency of antigen-specific IFN- γ -producing cells was determined by an ELISA spot technique²⁸ using the complementary IFN- γ mAbs. Frequency of antigen-induced spot-forming cells (SFC) is shown. Values are the mean $\pm$ s.e.m. from 5 female NOD mice, tested in triplicate cultures, with and without antigen. b, Detection of GAD-specific, IFN- γ -producing cells in the L-selectin⁺ (L-sel⁺) subpopulation of CD4⁺ cells. CD4⁺ cells were isolated from the spleens by panning on goat-anti-mouse immunoglobulin (Zymed) and on anti-CD8 (mAb 58.6-72)-coated plates. CD4⁺ cells were subfractionated using L-selectin-specific mAb MEL-14. CD4⁺ L-selectin⁺ and CD4⁺ L-selectin⁻ fractions were seeded in triplicate at 2×10^5 cells per well with irradiated (3,000 rad) spleen cells of 3-week-old NOD mice (5×10^5 cells per well) as a source of antigen-presenting cells. GAD-induced IFN- γ production was measured after 48 h by ELISA.

TABLE 1 Induced tolerance to GAD prevents the development of insulinitis and the spread of T-cell autoimmunity

Treatment	Insulinitis score	Spleen cell proliferation (SI $\pm$ s.e.m.)						
		β -galactosidase	GAD	GAD peptides			Hsp peptide	CPH
				No. 17	No. 34	No.35		
Uninjected	2.4 $\pm$ 0.2	1.0 $\pm$ 0.2	9.5 $\pm$ 2.1	4.8 $\pm$ 0.4	6.0 $\pm$ 0.1	2.9 $\pm$ 0.2	6.7 $\pm$ 1.0	ND
β -Galactosidase	2.6 $\pm$ 0.6	1.1 $\pm$ 0.1	<u>15.4 $\pm$ 1.8</u>	<u>5.1 $\pm$ 0.6</u>	<u>5.1 $\pm$ 0.6</u>	<u>4.0 $\pm$ 0.2</u>	<u>6.6 $\pm$ 0.5</u>	<u>11.5 $\pm$ 0.9</u>
GAD	0.1 $\pm$ 0.1	1.1 $\pm$ 0.03	1.6 $\pm$ 0.3	1.0 $\pm$ 0.05	1.2 $\pm$ 0.1	1.0 $\pm$ 0.1	1.2 $\pm$ 0.1	1.0 $\pm$ 0.02
Hsp-p	1.7 $\pm$ 0.4	1.1 $\pm$ 0.05	5.8 $\pm$ 0.2	4.5 $\pm$ 0.1	4.1 $\pm$ 0.3	4.2 $\pm$ 0.1	1.1 $\pm$ 0.04	4.4 $\pm$ 0.2
m-Hsp	1.8 $\pm$ 0.5	1.0 $\pm$ 0.1	4.2 $\pm$ 0.1	<u>3.9 $\pm$ 0.1</u>	<u>3.9 $\pm$ 0.2</u>	<u>3.4 $\pm$ 0.2</u>	1.0 $\pm$ 0.03	<u>4.3 $\pm$ 0.2</u>

Female NOD mice were injected intravenously at 3 weeks of age with 50 μ g GAD65, β -galactosidase, mycobacterial Hsp65 (m-Hsp) or 0.1 μ g immunodominant heat-shock peptide (Hsp-p) in PBS. At 12 weeks, mice were examined for insulinitis and splenic T-cell responses to IDDM-associated autoantigens. Pancreatic tissue sections were stained using immunoperoxidase for insulin and counterstained with haematoxylin. Insulinitis was scored blinded by examining 54 to 87 islets on 5 interrupted tissue sections from each pancreas. Severity of mononuclear cell infiltration was defined histologically: zero indicates lymphocytic infiltration; 1 indicates <25%; 2, 25–50%; 3, 50–75%; 4, >75%²³. Proliferative splenic T-cell responses were measured as described for Fig. 1. Values are stimulation index (SI) $\pm$ s.e.m. Significant T-cell responses are underlined; a dashed underline is used for a borderline response. See Fig. 3 legend for GAD peptide numbering system. CPH, carboxypeptidase H. ND, not determined at this time point. Background counts in all groups was typically around 2,000 c.p.m. $N=8$ for the GAD-treated group and $N=5$ for all other groups.

els similar findings in experimentally induced autoimmune disease^{16,17}. Apparently lymphokine secretion by the first wave of autoreactive T cells in the target organ induces loss of T-cell tolerance to additional antigens, resulting in a cascade of autoimmune responses^{18–20}. The activation of T cells reactive to additional β -cell antigens is likely to promote β -cell destruction (for example, Hsp-reactive CD4⁺ T-cell clones can induce IDDM^{11,12}).

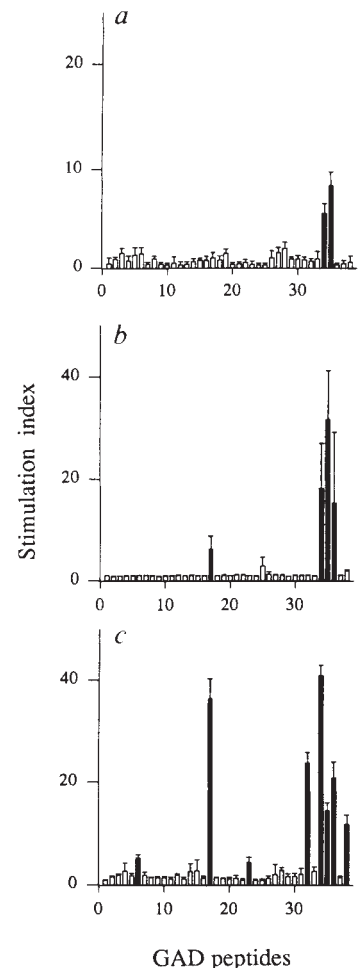
GAD-reactive T cells were the first to arise among the autoantigens tested (Fig. 1). But did the anti-GAD response itself develop through intermolecular spreading, after a T-cell response to an unidentified β -cell antigen? If the development of T-cell reactivity to GAD is a primary event in the pathogenesis of IDDM, the inactivation of GAD-reactive T-cells before their spontaneous priming should prevent the cascade of T-cell responses to other β -cell antigens, insulinitis and diabetes.

NOD mice were injected intravenously with GAD at 3 weeks of age, a treatment that causes unresponsiveness in antigen-reactive T cells²¹. Control groups received a similarly purified antigen (β -galactosidase), the immunodominant Hsp65 peptide, or mycobacterial heat-shock protein (which has been shown to vaccinate against murine IDDM¹¹). The mice were examined for autoantigen-reactive T cells and insulinitis at 12 weeks of age (Table 1), an age at which both reactive T cells and insulinitis are clearly established in untreated NOD mice. Seventy five per cent of the GAD-treated mice (but none of the controls), displayed no T-cell reactivity to GAD, indicating complete tolerization. The GAD-tolerized mice showed no reactivity to other β -cell antigens and were completely free of insulinitis (score zero). If there were another effector T-cell population in the islets, specific for an unknown β -cell antigen, that preceded the anti-GAD response, the release of cytokines by this population should have promoted T-cell responses to β -cell antigens and insulinitis^{19,20}. Twenty five per cent of the GAD-treated mice were not completely tolerized to GAD, as evidenced by weak residual GAD reactivity (stimulation index $\sim$ 3) and displayed very limited peri-insulinitis. In contrast, although tolerization to both of the heat-shock protein antigens was complete, these treatments reduced, but did not prevent, the development of T-cell responses to other β -cell antigens or insulinitis—as would be expected if a secondary element were removed from the amplificatory cascade.

In ongoing experiments examining the effects of GAD tolerization on diabetes incidence, all of the GAD-treated mice ($n=17$; now 37 weeks old) have normal glucose levels, whereas 70% of the mice receiving control antigens developed hyperglycaemia by 19 weeks of age ($n=20$). Five GAD-treated mice were examined at 30 weeks of age for insulinitis. Four pancreata were completely free of insulinitis (score 0.0) and one showed very limited peri-insulinitis (J.T. and D.L.K., unpublished results). Thus, GAD

FIG. 3 Intramolecular spreading of autoimmunity within the GAD molecule. Spleen cells were tested from 4-week-old (a), 5-week-old (b) and 7-week-old (c) NOD mice for proliferative responses to GAD65 peptides. A set of 38 peptides, each 20–23 amino acids long, span the entire human GAD65 (ref. 25) molecule with overlaps of 5 amino acids. These peptides are numbered successively from the N terminus. Peptides that triggered stimulation indices >3 are indicated as black bars. These peptides did not induce proliferation in T cells from NOD mice <3 or >16 weeks in age (paralleling reactivity to whole GAD), or from control (BALB/c $\times$ NOD)_{F1} mice (data not shown). Data are represented as the mean SI $\pm$ s.e.m. calculated from 3–6 mice tested individually in two separate experiments for each age group.

METHODS. Proliferation was assayed as described for Fig. 1. Peptides were present in cultures at 7 μ M and label was added during the last 16 h of a 5-day culture. Peptides were synthesized using standard F-moc chemistry and purified by reverse-phase HPLC (Advanced Chem-tech). The sequences of stimulatory peptides are: peptide 509–528 (no. 34), (IPPSRLRYLED–NEERMSRLSK); peptide 524–543 (no. 35), (SRLSKVAPVIKARMMMEYGT); and peptide 247–266 (no. 17), (NMYAMMIARFKMFPEVKEKG). Murine and human GAD65 share 95% amino-acid identity and are 98% conserved²⁵. Underlined amino acids are conservatively substituted in murine GAD65. In separate experiments, the murine forms of these peptides were tested and produced similar results.



tolerization can prevent the development of clinical diabetes in NOD mice.

Prevention of diabetes in NOD mice has also been reported following oral feeding of insulin²², immunization with complete Freund's adjuvant²³, heat-shock protein¹¹ and heat-shock protein-specific T cells¹². These treatments are thought to work by induction of active protective immune responses^{11,12,22,23} and have not been shown to prevent insulinitis. In contrast, we have shown that in the absence of T-cell reactivity to GAD, autoimmunity to β -cells does not develop, demonstrating the pathogenic significance of this response.

Our data qualify GAD as a key antigen in the induction of murine IDDM. The initial spontaneous anti-GAD response leads to the recruitment of additional β -cell antigen-reactive T cells. This cascade of autoimmune responses can be circumvented by inactivating GAD-reactive T cells. As a similar autoimmune progression may also occur in human IDDM²⁴ and in other T-cell-mediated autoimmune diseases (involving different autoantigens), these findings should be useful in the design of immunotherapies. □

Received 11 May; accepted 22 September 1993.

1. Castaño, L. & Eisenbarth G. S. A. *Rev. Immun.* **8**, 647–679 (1990).
2. Palmer, J. P. *Diabetes, Rev.* **1**, 104–115 (1993).
3. Kaufman, D. L., McGinnis, J. F., Krieger, N. R. & Tobin, A. J. *Science* **232**, 1138–1140 (1986).
4. Erlander, M. G., Tillakaratne, N. J. K., Feldblum, S., Patel, N. & Tobin, A. J. *Neuron* **7**, 91–100 (1991).
5. Kaufman, D. L. & Tobin, A. J. *Trends pharmac. Sci.* **14**, 107–109 (1993).
6. Baekkeskov, S. et al. *Nature* **298**, 167–169 (1982).

7. Baekkeskov, S. et al. *Nature* **347**, 151–156 (1990).
8. Atkinson, M. A., Maclaren, N. K., Scharp, D. W., Lacy, P. E. & Riley, W. J. *Lancet* **335**, 1357–1360 (1990).
9. Kaufman, D. L. et al. *J. clin. Invest.* **89**, 283–292 (1992).
10. Castaño, L., Russo, E., Zhou, L., Lipes, M. A. & Eisenbarth G. S. *J. clin. endocrin. Metab.* **73**, 1197–1201 (1991).
11. Elias, D., Markovits, D., Reshef, T., van der Zee, R. & Cohen, I. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1576–1580 (1990).
12. Elias, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3088–3091 (1991).
13. Croft, M., Duncan, D. & Swain, S. L. *J. exp. Med.* **176**, 1431–1437 (1992).
14. Bradley, L. M., Atkins, G. G. & Swain S. L. *J. Immun.* **148**, 324–331 (1992).
15. Ando, D. G., Clayton, J., Kono, D., Urban, J. L. & Sercarz, E. E. *Cell. Immun.* **124**, 132–143 (1989).
16. Lehmann, P. V., Forsthuber, T., Miller, A. & Sercarz, E. E. *Nature* **358**, 155–157 (1992).
17. Watanabe, R., Wege, H. & ter Meulen, V. *Nature* **305**, 150–153 (1983).
18. Lehmann, P. V., Sercarz, E. E., Forsthuber, T., Dayan, C. M. & Gammon, G. *Immun. Today* **14**, 203–208 (1993).
19. Sarvetnick, N. et al. *Nature* **346**, 844–847 (1990).
20. Heath, W. R. et al. *Nature* **359**, 547–549 (1992).
21. Gammon, G. & Sercarz, E. E. *Nature* **342**, 183–185 (1989).
22. Zhang, Z. J., Davidson, L., Eisenbarth, G. & Weiner, H. L. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10252–10256 (1991).
23. Qin, H.-Y., Sadelain, M. W. J., Hitchon, C., Lauzon, J. & Singh, B. J. *Immun.* **150**, 2072–2080 (1993).
24. Atkinson M. A. et al. *Lancet* **339**, 458–459 (1992).
25. Bu, D. F. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2115–2119 (1992).
26. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. & Stüber, D. *Bio/Technology* **6**, 1321–1325 (1988).
27. Macy, E., Kemeny, M. & Saxon, A. *FASEB J.* **3**, 3003–3009 (1988).
28. Taguchi, T. et al. *J. Immun.* **145**, 68–77 (1990).
29. Lee, D. S., Tian, J., Phan, T. & Kaufman, D. L. *Biochim. biophys. Acta* **1216**, 157–160 (1993).

ACKNOWLEDGEMENTS. We thank C. Evans and D. Keith for help and advice; C. Anderson for tissue preparation; M. Kronenberg, Z. Nagy, M. Cohn, N. Maclaren, M. S. Sy and D. Kaplan for comments on the manuscript; and D. Newman, T. Phan, A. Peck, N. Kim, W. Chen, K. Jensen, N. Khodadadi and V. Arroyo for their assistance. This work was supported by grants from the American Diabetes Association (D.L.K.), the NIH (M.C.S., A.J.T., E.E.S.), the Juvenile Diabetes Foundation (M.C.S., A.J.T., E.E.S.) and the Diabetes Research and Education Foundation (M.A.A.). P.V.L. was a Fellow of the Jaye Haddad/Concern Foundation and T.F. was an NIH Postdoctoral Trainee.

Immune response to glutamic acid decarboxylase correlates with insulinitis in non-obese diabetic mice

Roland Tisch, Xiao-Dong Yang, Steven M. Singer, Roland S. Liblau, Lars Fugger & Hugh O. McDevitt*

Departments of Microbiology and Immunology, and *Medicine, Stanford University Medical Center, Stanford, California 94305, USA

KNOWING the autoantigen target(s) in an organ-specific autoimmune disease is essential to understanding its pathogenesis. Insulin-dependent diabetes mellitus (IDDM) is an autoimmune disease characterized by lymphocytic infiltration of the islets of Langerhans (insulitis) and destruction of insulin-secreting pancreatic β -cells¹. Several β -cell proteins have been identified as autoantigens, but their importance in the diabetogenic process is not known². The non-obese diabetic (NOD) mouse is a murine model for spontaneous IDDM³. Here we determine the temporal sequence of T-cell and antibody responses in NOD mice to a panel of five murine β -cell antigens and find that antibody and T-cell responses specific for the two isoforms of glutamic acid decarboxylase (GAD) are first detected in 4-week-old NOD mice. This GAD-specific reactivity coincides with the earliest detectable response to an islet extract, and with the onset of insulitis. Furthermore, NOD mice receiving intrathymic injections of GAD65 exhibit markedly reduced T-cell proliferative responses to GAD and to the rest of the panel, in addition to remaining free of diabetes. These results indicate that the spontaneous response to β -cell antigens arises very early in life and that the anti-GAD immune response has a critical role in the disease process during this period.

We have cloned and expressed a panel of candidate murine β -cell autoantigens. This panel consists of: (1) the two isoforms

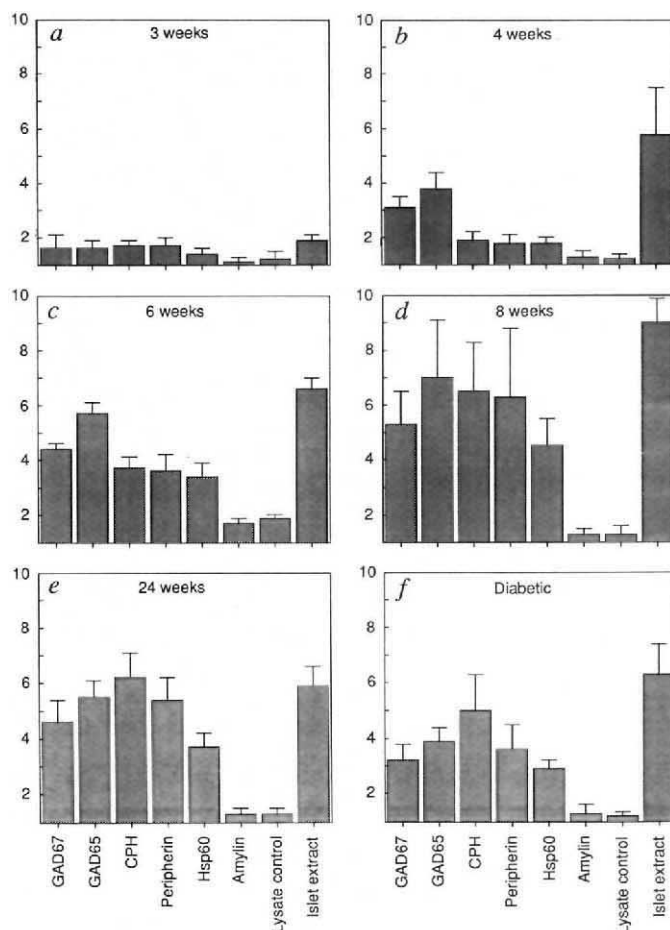
of GAD, GAD65 and GAD67; (2) peripherin; (3) carboxypeptidase H; and (4) heat-shock protein 60 (Hsp60). (GAD, an enzyme required for γ -aminobutyric acid synthesis, is recognized by T cells and antibodies in preclinical and recent-onset diabetics^{4–6}; peripherin is a cytoskeletal protein which is recognized by antibodies in the serum of NOD mice⁷; carboxypeptidase H, expressed in β -cell secretory granules, is recognized by antibodies in the serum of preclinical diabetics⁸; and Hsp60 stimulates T-cell and antibody responses in NOD mice⁹.) It is not known, however, whether any of these autoantigens has a critical role in the initial events in the diabetogenic process.

In NOD mice, β -cell destruction is T-cell-dependent^{10–13}. We therefore examined T-cell responses against the panel of β -cell antigens in unimmunized NOD mice. T-cell proliferation in response to an islet extract was first detected in splenic cell cultures established from 4-week-old females (Fig. 1*b*). There was a concomitant and significant response to GAD65 and GAD67. T-cell responses for peripherin, carboxypeptidase H and Hsp60 could first be seen in cultures established from 6-week-old NOD females (Fig. 1*c*). By 8 weeks of age, NOD mice showed enhanced responses to the entire panel (Fig. 1*d*). Furthermore, diabetic and 24-week-old non-diabetic NOD females also exhibited responses to the panel (Fig. 1*e,f*). In contrast, no response at any age could be detected with amylin, a 37-amino-acid polypeptide found in the insulin secretory granules of β -cells¹⁴ which binds I-A^{g7} (data not shown). This suggests that not all β -cell antigens become targets for the diabetogenic response. Finally, the response to the panel of β -cell antigens and to the islet extract can be completely blocked in cultures established from 8-week-old NOD female mice by addition of the anti-CD4 antibody GK1.5 (data not shown). Similar treatment with the anti-CD8 antibody 53.7.3 had no effect on the responses. Therefore, we are measuring a CD4⁺ T-cell proliferative response.

To examine further the *in vivo* effects of these T-cell responses, we measured antibodies specific for the panel of β -cell antigens. No response to the panel could be detected in sera from animals of 3 weeks old using an enzyme-linked immunoassay (Fig. 2),

FIG. 1 Spontaneous T-cell proliferation to β -cell antigens is age-dependent. Spleen cells from NOD/Mcd female mice at a, 3 weeks old; b, 4 weeks old; c, 6 weeks old; d, 8 weeks old; e, 24 weeks old, non-diabetic; and f, 20 weeks old, diabetic, were tested for their ability to respond to the panel of β -cell and control proteins. T-cell proliferation is expressed as stimulation index (mean c.p.m. of response to antigen divided by mean c.p.m. with medium only) on the y axis. The following range of c.p.m. were obtained in medium controls for the six individual mice tested in each group: 3 weeks old, 1,676–3,237; 4 weeks old, 1,992–3,122; 6 weeks old, 3,538–4,442; 8 weeks old, 1,447–3,424; 24 weeks old, non-diabetic, 1,994–4,160; 20 weeks old, diabetic, 1,997–4,383.

METHODS. Spleen cells prepared from individual mice were incubated in triplicate at $2.5 \times 10^6 \text{ ml}^{-1}$ for 72 h in 0.2 ml culture medium in the presence of $20 \mu\text{g ml}^{-1}$ antigen, a concentration giving optimal response. Proliferation was measured by the incorporation of [^3H]thymidine during the final 16 h of incubation. The standard error of the mean for each triplicate was less than 10%. Cell culture medium consisted of RPMI 1640, $5 \times 10^{-5} \text{ M}$ 2-mercaptoethanol and 1% Nutridoma-SP (Boehringer Mannheim). The cDNAs encoding murine GAD65, GAD67, peripherin, carboxypeptidase H (CPH) and Hsp60 were cloned and engineered to contain six histidine residues on the C terminus of each protein. Recombinant GAD65, GAD67, peripherin and CPH proteins were expressed in a baculovirus system¹⁹ and affinity-purified using a Ni^{2+} -conjugated resin (Invitrogen). Hsp60 produced in an *E. coli* expression system was similarly purified. Details of the cDNA cloning and protein purification of the antigens will be described elsewhere (R.T., manuscript in preparation). The lysate control was prepared from lysed Sf9 cells infected with wild-type baculovirus. The lysate was applied to the Ni^{2+} -resin and protein that bound to the resin was eluted and used for subsequent experiments. Diabetes-associated rat amylin, identical to murine amylin¹⁴, was from BaChem, California. Pancreatic islets were purified on a Ficoll gradient and β -cells sonicated to prepare the extract.



but sera from 4-week-old NOD mice showed a significant response to GAD65 and GAD67 (Fig. 2a, b), in parallel with the GAD-specific cellular responses already described. An enhanced response to GAD65 and GAD67 was observed in 6-week-old NOD females. Antibodies specific for carboxypeptidase H and peripherin could also be detected at this time (Fig. 2c, d). A response to Hsp60 was first observed in 8-week-old NOD females (Fig. 2e). Both diabetic and 24-week-old non-diabetic females continued to respond to the entire panel. No antibody response to the panel was seen in B10.GD mice, a non-diabetic strain which, like NOD, expresses H-2 K^d and D^b molecules, but differs from NOD by the expression of I-A^d.

To ensure that the T-cell responses observed in NOD mice are specific to this strain, we determined whether B10.GD, NOD, PD-A and NOR mice could also respond to the panel. NOD.PD-A is a NOD transgenic line expressing an I-A^g β -chain gene in which codons for histidine and serine at positions 56 and 57 have been mutated to proline and aspartic acid, respectively (S.M.S., manuscript in preparation). These mice are diabetes-free and develop pancreatic infiltration surrounding the islets (peri-insulinitis) infrequently. NOR mice are a diabetes-free strain

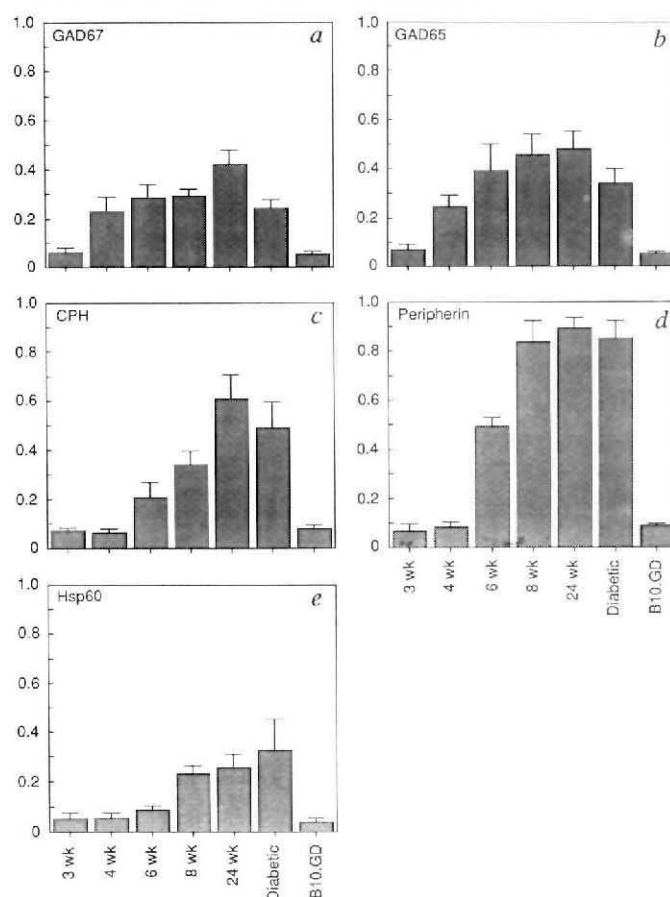


FIG. 2 Detection of antibodies specific for β -cell antigens is age-dependent. Three individual NOD/Mcd female mice of a specific age and 16-week-old B10.GD female mice were assessed for antibody responses to: a, GAD67; b, GAD 65; c, carboxypeptidase H; d, peripherin; and e, Hsp60 using an ELISA. Antibody detection is represented as units of optical density at 410 nm (y axis).

METHODS. 96-well plates (Dynatech) were coated with $10 \mu\text{g ml}^{-1}$ of antigen in 0.1 M NaHCO_3 , pH 8.5, overnight at 4°C . Wells were blocked with 0.1% gelatin in Tris-buffered saline, pH 8.0, and sera added at a dilution of 1/100. Bound antibody was detected with biotin-conjugated goat anti-mouse immunoglobulin and avidin alkaline phosphatase (Sigma). Reactions were developed with nitrophenylphosphate (Sigma) for 30 min and absorbance at 410 nm was read using an MR700 plate reader (Dynatech). The standard error of the mean for each duplicate was less than 10%.

with a major histocompatibility complex (MHC) identical to NOD, which exhibits peri-insulitis but has a very low frequency of infiltration directly within the islets (intra-insulitis)¹⁵. NOR mice differ genetically from NOD mice at parts of chromosomes 2, 4, 11 and 12. No significant proliferative response to the β -cell antigens was detected in the B10.GD cultures (Fig. 3). Similarly, the NOD.PD-A transgenic line was unresponsive to the panel. When we examined cultures established from 16-week-old NOR mice, a substantial response to the islet extract could be seen (Fig. 3f). Furthermore, a minimal yet statistically significant ($P < 0.001$) response, relative to the responses observed in B10.GD mice, could be detected in NOR mice to GAD65, GAD67, carboxypeptidase H and peripherin, but not to Hsp60. The reduced response seen for these five antigens in NOR relative to NOD appears to reflect the low levels of intra-insulitis seen in these mice. These findings support the idea that the panel of β -cell antigens is primarily involved in events associated with intra-insulitis. Moreover, T cells reactive to the panel of β -cell antigens are specific to strains that develop islet cell autoimmunity.

The early cellular and antibody responses to GAD suggest that this β -cell autoantigen may be critical in the initial stages of the diabetogenic process, so we injected GAD65 into the thymuses of NOD mice to tolerize the animals to this antigen. It has been shown¹⁶ that NOD mice injected with β -cells in the thymus at 4 weeks old become tolerant to islet cells and remain

TABLE 1 Intrathymic injection of GAD65 reduces the frequency of insulitis

Treatment	No. of mice	Severity of insulitis		
		—	+	++
Sf9 Extract	2	1/19 (5%)	5/19 (26%)	13/19 (69%)
GAD65	6	51/71 (72%)	12/71 (17%)	8/71 (11%)

Formalin-fixed pancreata from non-diabetic mice intrathymically injected with either Sf9 protein extract or GAD65 were sectioned and stained with haematoxylin and eosin. The severity of insulitis seen in the islets of the Sf9 protein extract injected mice ($n=19$) and the GAD65-injected mice ($n=71$) was assessed by the following criteria: —, islets free of insulitis; +, islets exhibiting peri-insulitis; ++, islets exhibiting intra-insulitis.

diabetes-free. It can be seen from Fig. 4 that the time of onset and the frequency of diabetes in NOD females injected intrathymically at 3 weeks old with GAD65 differ significantly from the group receiving an injection of an Sf9 cell extract. Furthermore, diabetic animals in the GAD65-injected group continue to respond to GAD65, suggesting that these mice are not tolerized by the procedure (Fig. 5b). In contrast, the animals injected with GAD65 that remain diabetes-free show a significantly reduced T-cell proliferative response to GAD65 relative to GAD65-injected diabetic mice (two-sided Student t -test, $P < 10^{-3}$) and to the control group ($P < 10^{-4}$) (Fig. 5a, b). The T-cell response to GAD67 and to the remainder of the panel is also significantly reduced in non-diabetic mice relative to the control group. In addition, histological examination of the pancreata from the non-diabetic, GAD65-injected mice revealed a significant reduction in intra-insulitis (Table 1). Moreover, this protection from disease seems to be specific for GAD65. NOD mice injected intrathymically with peripherin became diabetic with the same frequency as animals injected with Sf9 protein extract (Fig. 4). Analysis of T-cell proliferation in the peripherin-injected mice indicated that the diabetic animals exhibited a significantly reduced response to peripherin ($P < 10^{-4}$), but continued to respond to the rest of the panel compared with mice injected with Sf9 protein extract (Fig. 5c). A T-cell response

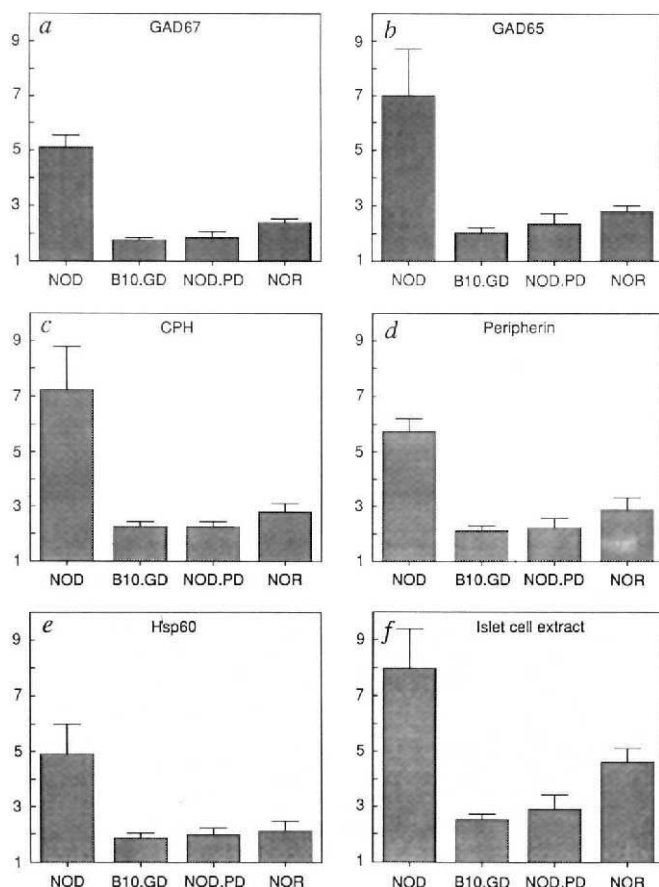


FIG. 3 Spontaneous T-cell proliferation to β -cell antigens is strain-specific. 16-week-old female NOD/Mcd, B10.GD, NOD.PD and NOR/Lt mice were individually assessed for a T-cell proliferative response to a, GAD67; b, GAD65; c, carboxypeptidase H; d, peripherin; e, Hsp60; and f, islet extract. T-cell proliferation is expressed as a stimulation index on the y axis. The following ranges of c.p.m. were obtained in medium controls for the six individual mice in each group: NOD/Mcd, 1,948–4,074; B10.GD, 6,637–8,924; NOD.PD-A, 3,828–4,936; NOR/Lt, 3,421–6,338. Methods are described in Fig. 1 legend. The standard error of the mean for each triplicate was less than 10%.

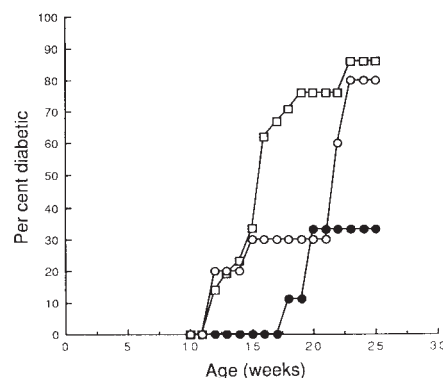


FIG. 4 Intrathymic injection of GAD65 has a protective effect from diabetes. NOD mice receiving intrathymic injections of GAD65 ($n=9$) (filled circles) show a delayed onset ($P = 0.001$; Mann-Whitney test) and reduced frequency ($P < 0.01$; one-sided Fisher's exact test) of diabetes compared with mice intrathymically injected with an Sf9 protein extract ($n=21$) (open squares). In contrast, NOD mice receiving an intrathymic injection of peripherin ($n=10$) (open circles) show no significant delay in disease onset and a frequency of diabetes similar to that seen in the Sf9 protein extract-injected animals. The data presented for the Sf9 protein extract-injected mice represent two separate experiments. METHODS. Three-week-old NOD female mice had 10 μ g GAD65 or peripherin or 20 μ g of an Sf9 protein extract injected into both lobes of the thymus. The Sf9 protein extract was prepared as described in Fig. 1 legend. Animals were checked weekly for hyperglycaemia and were considered to be diabetic if they had glycaemia levels >200 mg dl^{-1} .

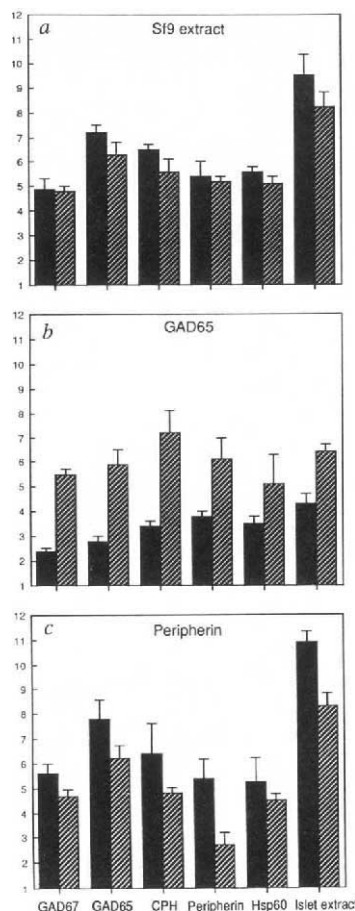


FIG. 5 Non-diabetic mice injected intrathymically with GAD65 lack T-cell responses to GAD and to the panel of antigens, whereas tolerance induction to peripherin has no such effect. Mice intrathymically injected with an Sf9 protein extract, GAD65 or peripherin were individually assessed for a T-cell proliferative response (a, b and c, respectively) to the panel of antigens. T-cell proliferation is expressed as a stimulation index on the y axis. The range of c.p.m. obtained in medium controls for 16 individual mice injected with the control extract, consisting of 3 non-diabetic (solid bars; $n=3$) and 13 diabetic (hatched bars; $n=13$) animals, the 9 individual mice injected with GAD65, consisting of 6 non-diabetic ($n=6$) and 3 diabetic ($n=3$) animals, and the 10 individual mice injected with peripherin, consisting of 2 non-diabetic ($n=2$) and 8 diabetic animals was 2,684–4,757, 2,707–6,212 and 1,609–4,163, respectively. Methods are described in the legends to Figs 1 and 2. The standard error of the mean for each triplicate in the T-cell proliferation assay was $<10\%$.

to peripherin was observed in the two non-diabetic animals, suggesting that tolerance induction in these mice had failed. Interestingly, despite the lack of a T-cell proliferative response, the non-diabetic, GAD65-injected mice continued to show a humoral response to GAD and the other antigens in the panel that was similar to the control group (data not shown). These results show that intrathymic injection of GAD65 leads to suppression of a T-cell proliferative response to GAD and to the other antigens, resulting in a marked decrease in intra-insulinitis and subsequent protection from overt diabetes. In contrast, suppression of a T-cell proliferative response to peripherin has no apparent effect on disease frequency.

The diabetogenic process develops sequentially. Peri-insulinitis is first detected in NOD mice at 4 weeks of age. These lesions then progress to intra-insulinitis, eventually leading to β -cell destruction and diabetes. We have shown that these events correlate with a temporal progression of autoreactivity to specific β -

cell antigens, possibly reflecting a cascade of antigen recognition. NOR mice with peri-insulinitis and a minimal frequency of intra-insulinitis respond to the islet extract but only marginally to the panel of autoantigens. Only when there is a frequent progression to intra-insulinitis do we detect a strong T-cell response to GAD and to the rest of the panel (Fig. 3). In turn, progression to intra-insulinitis is effectively blocked by suppression of GAD auto-reactivity. A response to GAD appears to be necessary but not sufficient to allow the disease to progress to the diabetic state. This is apparent in 24-week-old non-diabetic NOD females which, despite their significant intra-insulinitis and proliferative responses to GAD, remain euglycaemic. It is possible that T-cell reactivity to other antigens not in this panel is required to complete destruction of β -cells, leading to diabetes.

Mice injected with GAD65 intrathymically remain diabetes-free and show a marked reduction in spontaneous T-cell proliferation to GAD and to the rest of the panel. Interestingly, an antibody response to the panel persists, suggesting that a form of 'split tolerance' may have been induced in these non-diabetic mice. This protective effect could result from the induction of $CD4^+$ Th2 T cells, which inhibit Th1 T-cell-mediated immunity but promote humoral immunity¹⁷. Indeed, there is a significant reduction in interferon- γ production, a product of Th1-type T cells, in antigen-stimulated cultures established from the GAD65-treated non-diabetic mice (data not shown). Alternatively, injecting GAD directly into the thymus may induce negative selection, thereby reducing the frequency of GAD-specific T cells in the periphery. This would prevent insulinitis and further destruction of β -cells and explain why the more distally observed T-cell proliferative responses to carboxypeptidase H, peripherin and Hsp60 are suppressed in the GAD-injected mice.

In humans, anti-GAD reactivity is found in a relatively high proportion of asymptomatic and recent diabetic patients^{6,18}. Similarly, we have shown that in NOD mice, both cellular and humoral responses to GAD are detectable at the same time as the initial signs of disease. Furthermore, this anti-GAD reactivity precedes the responses to any of the other autoantigens examined. Finally, intrathymic injection of GAD65 elicits protection from diabetes, whereas tolerance to peripherin, a β -cell antigen recognized later in the disease process, is insufficient to induce a similar form of protection. Together these observations strongly indicate that GAD plays a critical part in the initial events in IDDM, although the possibility that additional β -cell antigens also participate in the initial stages of the diabetogenic process cannot be ruled out. □

Received 14 May; accepted 5 October 1993.

1. Eisenbarth, G. S. *New Engl. J. Med.* **314**, 1360–1368 (1986).
2. Harrison, L. C. *Immun. Today* **13**, 348–352 (1992).
3. Makino, S. et al. *Exptl Anim.* **29**, 1 (1980).
4. Baekkeskov, S. et al. *Nature* **298**, 167–169 (1982).
5. Baekkeskov, S. et al. *Nature* **347**, 151–166 (1990).
6. Honeyman, M. C., Cram, D. S. & Harrison, L. C. *J. exp. Med.* **177**, 535–540 (1993).
7. Boitard, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 172–176 (1992).
8. Castano, L., Russo, E., Zhou, L., Lipes, M. A. & Eisenbarth, G. S. *J. clin. Endocrin. Metab.* **73**, 1197–1201 (1991).
9. Elias, D., Markovits, D., Reshef, T., van der Zee, R. & Cohen, I. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1576–80 (1990).
10. Shizuru, J., Taylor-Edwards, C., Banks, B. A., Gregory, A. K. & Fathman, C. G. *Science* **240**, 659–662 (1988).
11. Miller, B. J., Appel, M. C., O'Neil, J. J. & Wicker, L. S. *J. Immun.* **140**, 52–58 (1988).
12. Hayward, A. & Schreiber, M. J. *Immun.* **143**, 1555–1559 (1989).
13. Reich, E. P., Sherwin, R. S., Kanagawa, O. & Janeway, C. A. *Nature* **341**, 326–328 (1989).
14. Leffert, J. D., Newgard, C. B., Okamoto, H., Milburn, J. L. & Luskey, K. L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3127–3130 (1989).
15. Prochazka, M., Serreze, D. V., Frankel, W. N. & Leiter, E. H. *Diabetes* **41**, 98–106 (1992).
16. Gerling, I. C. et al. *Diabetes* **41**, 1672–1676 (1992).
17. Coffman, R. L. et al. *Immun. Rev.* **102**, 5–28 (1988).
18. Atkinson, M. A. et al. *Lancet* **339**, 458–459 (1992).
19. O'Reilly, D. R., Miller, L. K. & Luckow, V. A. *Baculovirus Expression Vectors: A Laboratory Manual* (Freeman, New York, 1992).

ACKNOWLEDGEMENTS. X.-D.Y. and S.M.S. contributed equally to this work. We thank P. Liblau, S. Patel and M. Sullivan for help and discussion, and E. Leiter for NOR/Lt mice. This work was supported by grants from the NIH. R.T., X.-D.Y., S.M.S., R.S.L. and L.F. were supported by fellowships from the MRC of Canada and the NIH, the American Diabetes Association, Howard Hughes Medical Institute, Juvenile Diabetes Foundation and the Alfred Benzon Foundation of Denmark, respectively.

B70 antigen is a second ligand for CTLA-4 and CD28

Miyuki Azuma*, Daisuke Ito*†, Hideo Yagita*,
Ko Okumura*, Joseph H. Phillips‡,
Lewis L. Lanier‡ & Chamorro Somoza‡

* Department of Immunology, Juntendo University School of Medicine, Hongo 2-1-1, Bunkyo-ku, Tokyo 113, Japan

† Department of Oral and Maxillo-facial Surgery, Faculty of Medicine, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113, Japan

‡ Department of Immunology, DNAX Research Institute of Molecular and Cellular Biology, 901 California Avenue, Palo Alto, California 94304, USA

THE membrane antigen B7/BB1 (refs 1, 2) is expressed on activated B cells, macrophages and dendritic cells, and binds to a counter-receptor, CD28, expressed on T lymphocytes and thymocytes³. Interaction between CD28 and B7 results in potent costimulation of T-cell activation initiated through the CD3/T-cell receptor complex^{4,5}. Discrepancies between results with anti-CD28 and anti-B7 antibodies have suggested the existence of a second ligand for CD28 and CTLA-4 (refs 3, 6–8). We have generated a monoclonal antibody, IT2, that reacts with a 70K glycoprotein (B70). B70 complementary DNA was cloned from a B-lymphoblastoid cell line library and encodes a new protein of the immunoglobulin superfamily with limited homology to B7. B70 is expressed on resting monocytes and dendritic cells and on activated, but not resting, T, NK and B lymphocytes. IT2 substantially inhibited the binding of a CTLA4-immunoglobulin fusion protein to human B-lymphoblastoid cell lines and, together with anti-B7 antibody, completely blocked CTLA-4 binding. Further IT2 efficiently inhibited primary allogeneic mixed lymphocyte responses. These findings indicate that B70 is a second ligand for CD28 and CTLA-4 and may play an important role for costimulation of T cells in a primary immune response.

Chimaeric fusion proteins containing the extracellular domain of CTLA-4 (ref. 23) and immunoglobulin γ bind B7 with a greater affinity than CD28 (ref. 24) and efficiently block T-cell-dependent immune responses *in vitro* and *in vivo*^{9–11}. Thus, CTLA4-immunoglobulin (CTLA4-Ig) fusion proteins can be used to detect cells bearing B7 and potentially other novel ligands.

A human NK-like leukaemia cell line, YT2C2, preferentially kills cell lines expressing B7 (ref. 6). To substantiate the existence of another ligand, we generated a monoclonal antibody, IT2, that blocked YT2C2-mediated cytotoxicity against JY B-lymphoblastoid cell line (B-LCL) cells in the presence of anti-B7 monoclonal antibody. As previously shown⁶, anti-CD28 antibody completely inhibits YT2C2 cytotoxicity against JY cells, whereas anti-B7 antibody only partially blocks cytotoxicity (Fig. 1a). When IT2 was combined with anti-B7 antibody, lysis of JY by YT2C2 was essentially abrogated (Fig. 1a). Because YT2C2 expresses CD28, but not CTLA-4 (as determined by analysis of messenger RNA expression; our unpublished observation), these data indicate that the antigen recognized by IT2, designated B70, functionally interacts with CD28 expressed on YT2C2 cells.

IT2 reacted with B-LCL JY cells but did not react with human B7-transfected murine P815 mastocytoma cells (Fig. 1b). IT2 also failed to react with murine L cells or human Jurkat leukaemia cells transfected with human B7 (not shown). IT2 substantially inhibited the binding of CTLA4-Ig to JY cells but did not affect the CTLA4-Ig binding to B7-P815 cells (Fig. 1b). When JY cells were pre-coated both with IT2 and with anti-B7 antibodies, binding of CTLA4-Ig was completely inhibited, indicating that B70 also binds to CTLA-4.

IT2 was used to clone and express the cDNA encoding the B70 polypeptide from a JY cDNA library. COS7 cells transfected with B70 cDNA react with IT2 and CTLA4-Ig, but not

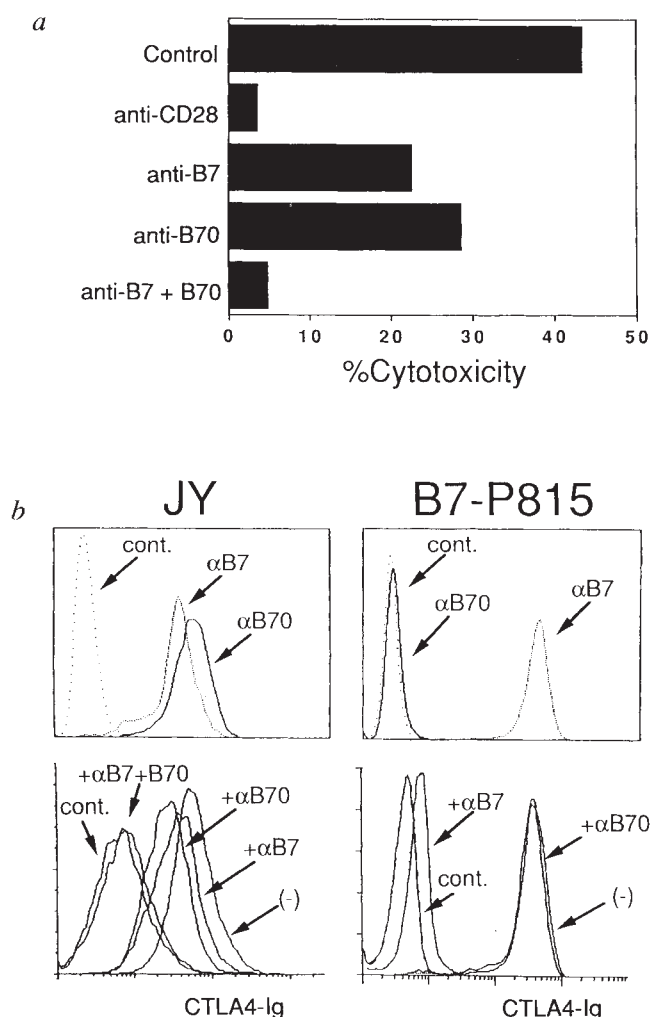


FIG. 1 IT2 (anti-B70) antibody. **a**, Inhibitory effect on YT2C2 lysis of JY cells. Cytotoxicity against human B-LCL cells was measured by a standard 4-h ⁵¹Cr-release assay at an effector:target ratio of 40:1. Control antibody, anti-CD28 (L293), anti-B7 (L307) and anti-B70 (IT2.1) antibodies were used at a final concentration of 5 μ g ml⁻¹ and were added at the start of the assay. Representative data are shown. Similar results were obtained from the experiments using several B7 and B70 double-positive B-LCL cell lines. **b**, JY and human B7-transfected murine P815 mastocytoma cells⁷ were stained with control immunoglobulin, anti-B7 or anti-B70, followed by FITC goat anti-mouse immunoglobulin. For CTLA4-Ig binding inhibition studies, JY or B7⁺ P815 were pre-coated with control murine immunoglobulin (—), murine anti-human B7, murine anti-human B70 or a combination of anti-B7 and anti-B70, followed by staining with CTLA4-Ig human-immunoglobulin fusion protein, and visualization with FITC-conjugated goat anti-human IgG. **METHODS**. IT2 antibodies (IT2.1, IgG1; IT2.2, IgG2b) were generated by fusing P3U1 myeloma cells with splenocytes from BALB/c mice immunized with JY B-LCL. Hybridoma supernatants were screened in a 4-h ⁵¹Cr release assay for the ability to block the killing of JY cells by YT2C2 in the presence of saturating concentrations of anti-B7 antibody (L307)⁶. In the CTLA4-Ig binding experiments, JY or B7⁺ P815 were stained with saturating amounts of irrelevant isotype-matched control antibody, L307 (IgG1), IT2.1 (IgG1) or mixture of L307 and IT2.1 for 20 min before the addition of CTLA4-Ig. Cells were washed and then stained with FITC-conjugated goat anti-human IgG (CalTag). Immunofluorescence and flow cytometry were done using a FACScan (Becton Dickinson Immunocytometry Systems) as described¹⁹. Data are displayed as histograms. The x-axis represents fluorescence and the y-axis the relative cell number. Histograms from cells stained with control antibody (nearest the ordinate) are superimposed over the histogram of cells stained with anti-B7, anti-B70 or CTLA4-Ig, as indicated. Anti-B7 antibody (L307) and other anti-human leukocyte antigen antibodies were provided by Becton Dickinson. A mouse CTLA-4-human IgG fusion protein was provided by P. Lane (Basel Institute).

anti-B7 antibody BB1 or L307 (Fig. 2a). Nucleotide sequencing revealed that the B70 cDNA encodes a 323-amino-acid polypeptide composed of an N-terminal signal peptide, a 223-amino-acid extracellular segment, a 23-amino-acid transmembrane domain, and a 61-amino-acid cytoplasmic domain (Fig. 2b). The extracellular domain contains eight sites for *N*-linked glycosylation and the intracellular region has three potential sites for protein kinase C-dependent phosphorylation. The extracellular portion of B70, like that of B7, contains a V-set and a C2-set immunoglobulin-like domain. Despite their overall similar structure, B70 has only limited amino-acid homology to human and mouse B7 (Fig. 2c), which was unexpected given that they share common and highly related receptors, CD28 and CTLA-4. Overall, human B70 has 23% identity with human B7 and 27% identity with mouse B7.

A single glycoprotein of roughly M_r 70K was immunoprecipitated from 125 I-labelled NK cell clones in both reducing (Fig. 3) and non-reducing (not shown) conditions. Consistent with an earlier report², anti-B7 antibody immunoprecipitated a smaller protein of ~55–60K. Affinity-purified B7 and B70 glycoproteins

were treated with *N*-glycanase and analysed by SDS-PAGE. A ~40K polypeptide was revealed after deglycosylation of B70, compared to the ~30K deglycosylated B7 polypeptide (Fig. 3). CTLA4-Ig immunoprecipitated two distinct molecules of ~60K and ~70K from biotin-labelled JY lysates and pre-clearing of lysates with anti-B7 or anti-B70 removed the 60K and 70K proteins, respectively, suggesting that B7 and B70 were independent molecules (not shown). Affinity-purified, biotin-labelled B70 glycoprotein was precipitated with CTLA4-Ig, but not with anti-B7 antibody, demonstrating that B70 directly interacts with CTLA-4 (not shown).

Flow cytometric analysis revealed that B70 was highly expressed on human B-LCL (JY, WA, NC37, RPMI-8866) and Burkitt's B lymphoma cell lines (Daudi, Raji), whereas resting peripheral blood B cells (Fig. 4A, b) and pre-B leukaemia cell lines (NALM6, SMS-SB, 697) (not shown) expressed only low levels of B7 or B70. B70, like B7 (refs 13–16), is induced after lymphocyte activation. Crosslinking of CD40 on B cells resulted in strong expression of B70 and B7 (Fig. 4A, d). Lipopolysac-

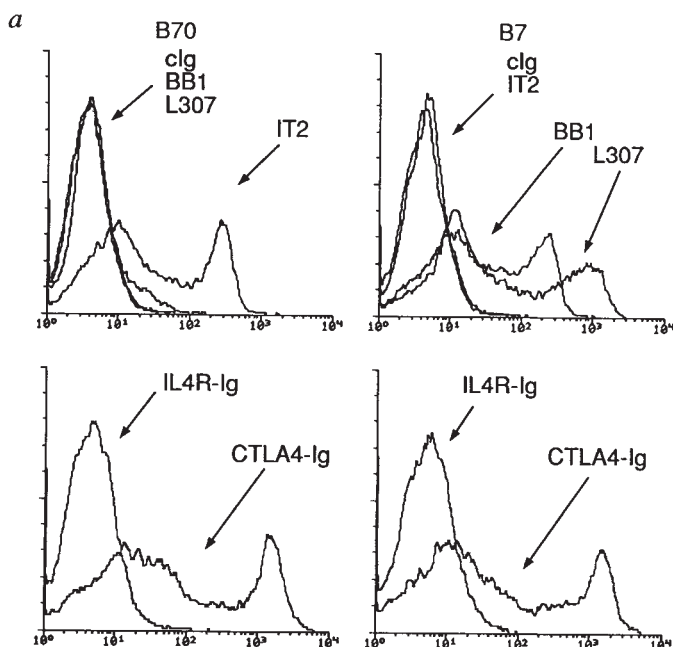
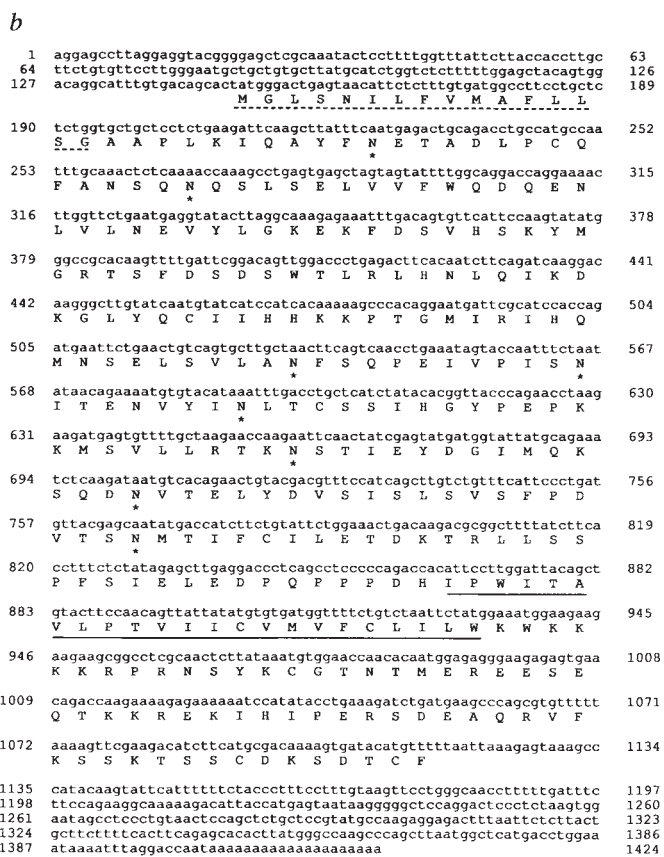


FIG. 2 Cloning and expression of B70 cDNA. a, COS7 cells transfected with human B70 or B7 cDNA and stained with: (upper panels) control immunoglobulin, anti-B70 antibody IT2, or anti-B7 antibody L307 and BB1, followed by phycoerythrin (PE)-conjugated anti-mouse immunoglobulin; (lower panels) murine IL-4 receptor-human immunoglobulin fusion protein (control) or murine CTLA4-human immunoglobulin fusion protein, followed by FITC-conjugated goat anti-human immunoglobulin. b, Nucleotide and predicted amino-acid sequence of human B70. Signal peptide is indicated by a broken underline, transmembrane is indicated by a solid underline, and sites for *N*-linked glycosylation are indicated by asterisks. c, Alignment of the extracellular domains of human B7, mouse B7, and human B70 protein sequences. Identical residues are shown in reverse lettering and conservative substitutions are boxed. METHODS. COS7 cells were transfected using the DEAE-dextran method with a JY B-LCL cDNA library in the pJFE14 expression vector, and stained with anti-B70 antibody and PE-conjugated goat anti-mouse immunoglobulin²⁰. Positive cells were isolated by flow cytometry and plasmids recovered by Hirt precipitation and bacterial transformation. After four rounds of enrichment, individual plasmids encoding B70 antigen were isolated and the cDNA insert was subcloned into Bluescript and sequenced by the dideoxy chain termination method. cDNA sequence of a 1.5 kilobase (kb) clone encoding B70 is shown.



c

hB7
mB7
hB70

GLSHFCSGV.IH[V]TKEV[K]E[V]A[LS]G[H].NVSVEELAQTR[I]Y[W]CKEKK[V]L.T.MM8[DM]
RLSQVSSDVEQLSKSVKQK[V]LLP[RY].NSPHEDESDRI[Y]WQKHDKV[LS].VIAQKL
AAPL[K]QAYFNE[T]A[LP]Q[FA]NSQNSLSL[V]F[W]QDQENL[V]LNEVYL[KE]

hB7
mB7
hB70

N...[I]WPEY[K]NRT[IFD]TNNL[S]I[V]LALRPSDE[ET]VE[V]VLYKEKDAFKRE[LA]EVL[
K...[V]WPEY[K]NRT[LY]NT.TYSL[I]LGLVLSDRCT[YS]CVVQKKERTGYEVK[L]ALV[L]
KFD[S]VHKKY[M]ORT[S]F[S]DSW[TL]L[H]NL[Q]IKDK[LY]O[C]I[H]HKKPTGM[IR]T[OM]NSEL

hB7
mB7
hB70

S[V]KAD[P]T[PS]SDFE[IP]T[SW]I.RR[T]CSTSGG[F]EP.HLSW[E...][Q]EELNAINTTV
S[V]KAD[P]T[PS]N[TS]GWP[S]ADT.KR[IT]CFASG[F]K[P].RFSW[E...][N]GRELPGINTI
S[V]LAN[S]OPE[V]PIS[IN]ENVY[IN]L[TC]SS[H]G[V]EPKKMSV[L]RT[N]STIEYDGMOK

hB7
mB7
hB70

SODPETELYA[V]ESK[DF...NM]T[TH]N[FM]C[L]KYGH[L]V[N]QTFNWN[T]TKQ[E]H...FPDN
SODPESELYT[IS]SQ[DF...NT]T[TH]N[FM]C[L]KYGH[L]V[N]QTFNWN[T]TKQ[E]H...FPDN
SODNVTLEYD[VS]IS[SV]SFFDVT[SN]M[T]F[C]IL[ET]DKT[HL]LSSP.FSIELE[DP]PPPPDH

FIG. 3 Immunoprecipitation of IT2 antigen. B7 and B70 were immunoprecipitated from a 125 I-labelled NK cell clone and were treated with *N*-glycanase to remove complex *N*-linked carbohydrates, as indicated.

METHODS. Viable NK cell clones were surface labelled with 125 I, lysed in 1% NP-40/Tris-buffered saline, and antigens were immunoprecipitated using Pansorbin (Calbiochem) coated with rabbit anti-mouse immunoglobulin and IT2.2 or L307, as described previously²¹. Samples were treated with 15 U ml⁻¹ recombinant *N*-glycanase F (Boehringer-Mannheim) for 18 h (ref. 21) and then analysed by SDS-PAGE using 10–20% gradient gels. Enzyme concentration and digestion time were optimized to ensure complete removal of *N*-linked oligosaccharides. Viable JY cells were surface biotinylated, lysed in 1% NP-40/Tris-buffered saline, and immunoprecipitated using control immunoglobulin, anti-B70 (IT2.2), anti-B7 (L307) or CTLA4-Ig (ref. 22). For sequential immunoprecipitation, biotin-labelled JY lysates were first precleared with L307 or IT2.2 antibodies and then immunoprecipitated with the individual antibodies. B70 antigen was affinity-purified from biotin-labelled JY cells using IT2.2 conjugated to CNBr-activated Sepharose, antigen was eluted with 50 mM diethylamine (pH 12), neutralized with 1 M Tris, pH 7.6, diluted in 1% NP-40/Tris-buffered saline containing 10 mg ml⁻¹ BSA and protease inhibitors, and re-immunoprecipitated with control immunoglobulin, CTLA4-Ig, IT2.2 and L307 antibodies.

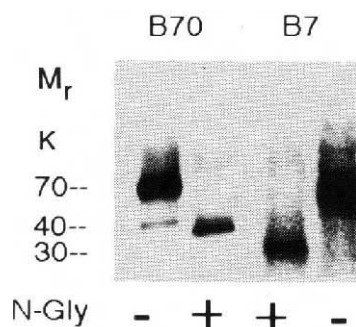
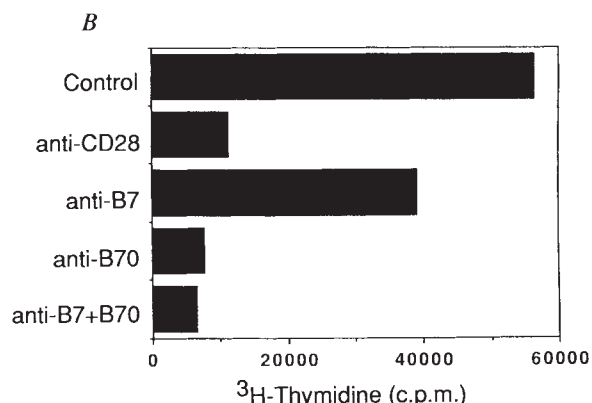
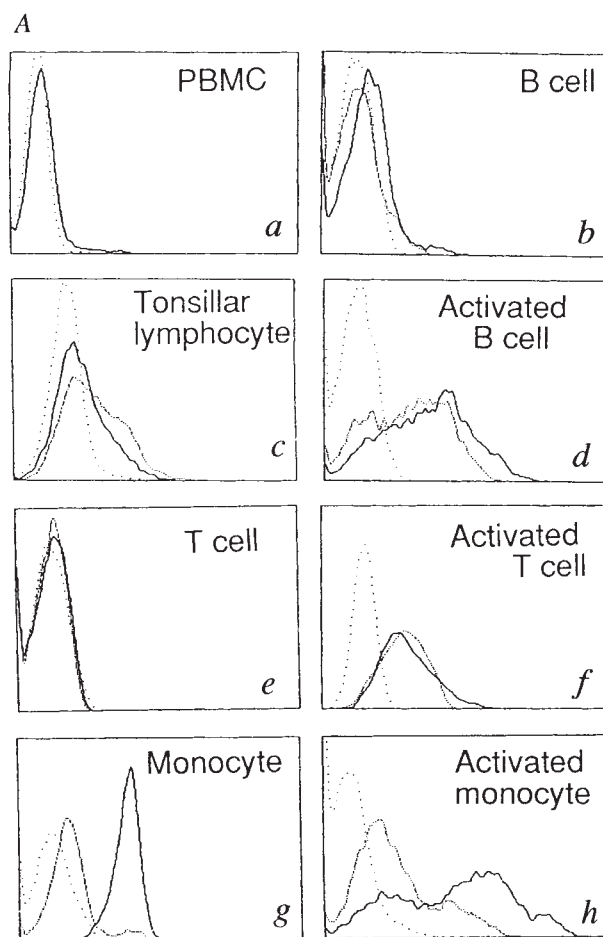


FIG. 4 A, Expression of B70 and B7 on human leukocytes. Expression of B70 (solid lines) and B7 (closely dashed lines) on: a, freshly isolated peripheral blood lymphocytes; b, freshly isolated CD19⁺ B cells; c, freshly isolated tonsillar lymphocytes resected with chronic tonsillitis; d, anti-CD40 activated B cells; e, freshly isolated CD3⁺ T cells; f, anti-CD3 and IL-2-activated peripheral blood T cells; g, freshly isolated CD14⁺ monocytes; h, γ -interferon-activated monocytes. B, IT2 efficiently inhibits primary allogeneic MLR. PBMC were isolated by density gradient centrifugation using Ficoll-Paque. Monocytes and B cells were depleted by plastic adherence and passage through nylon wool columns⁷. Non-adherent lymphocytes as responder cells (10⁵ cells per well) were cocultured with 2 $\times$ 10⁴ irradiated allogeneic PBMC in the presence of control antibody, anti-B70 (IT2.1), anti-B7 (L307), a mixture of anti-B7 and -B70, or anti-CD28 (L293). Cells were cultured in 96-well flat-bottomed plates for 7 days, labelled for the final 18 h with 3 H-thymidine, and collected⁷. Incorporated radioactivity was measured in a β -scintillation counter. Each value is the mean of triplicate wells. Representative data are shown from five experiments using different blood donors. Each antibody was added at a final concentration of 5 μ g ml⁻¹ at the initiation of the assay.

METHODS. A, PBMC were isolated from healthy human volunteers by using Ficoll-Paque density gradients. For activating B cells, peripheral blood B cells were isolated with anti-CD19 coated Dynabeads and DetachaBeads (Dyna), according to the manufacturer's recommendation, and were cultured with anti-CD40 (100 ng ml⁻¹) antibody and irradiated CDw32-transfected L cells in the presence of rIL-4 (10 ng ml⁻¹) as described¹⁶. Anti-CD40 antibody, rIL-4 and CD32-L cells were from H. Kikutani, H. Spits and K. Moore, respectively. For activating T cells, freshly isolated PBMC were cultured in medium containing 100 U ml⁻¹ rIL-2 (Shionogi Pharmaceutical, Japan) for 10 days in tissue culture flasks coated with purified anti-CD3 antibody, as described¹⁵. For activating monocytes, freshly isolated PBMC were cultured with 500 U ml⁻¹ rIFN- γ (Shionogi) for 24 h. Freshly isolated and cultured cells were stained with biotin-conjugated anti-B70 (IT2.2) or biotin-conjugated anti-B7 (L307) together with FITC anti-CD14 (Leu M3), FITC anti-CD19 (Leu 12) or FITC anti-CD3 (Leu 4), followed by PE-streptavidin. For two-colour immunofluorescence, an electronic gate was set on FITC-stained CD3-, CD19- or CD14-positive cells to identify T, B and monocytes, respectively, and PE-labelled anti-B70 and anti-B7 fluorescence was displayed. Histograms from cells stained with control antibody (dotted lines, nearest the ordinate) are superimposed over the histograms of cells stained with anti-B7 and anti-B70.

haride (LPS)- or pokeweed mitogen (PWM)-activated B cells (not shown) and freshly isolated tonsillar B cells resected with chronic tonsillitis (Fig. 4A,c) expressed significant levels of B70, as well as B7. Similarly, although peripheral blood T cells (Fig. 4A,e) were negative, both B7 and B70 were induced after stimulation for 10 days with immobilized anti-CD3 antibody and recombinant interleukin-2 (rIL-2; Fig. 4A,f). CD4⁺ T-cell clones, CD8⁺ T-cell clones and NK cell clones expressed B70 (not shown), as well as B7 (ref. 15), whereas T leukaemia cell lines (Jurkat, Molt 4, PEER, CEM, HUT-78) did not express these antigens (not shown). Although B7 and B70 are apparently coexpressed by normal, activated T and B lymphocytes, their expression differs in myeloid cells. Freshly isolated peripheral blood monocytes constitutively express significant levels of B70, with only much lower amounts of B7 (Fig. 4A,g). After incubation for 24 h with γ -interferon, B70 on monocytes was greatly enhanced and B7 was induced (Fig. 4A,h). Peripheral



blood dendritic cells expressed B7 (ref. 17) and high levels of B70 (not shown).

B7 costimulation is required for the efficient generation of a primary mixed lymphocyte response (MLR) directed against alloantigens. Anti-B7 antibodies partially inhibit allogeneic MLR^{5,17}, and transfection of B7-negative B-cell lines with B7 cDNA substantially augments their ability to stimulate primary T-cell responses^{7,18}. To examine the function of B70, freshly isolated nonadherent peripheral blood lymphocytes (PBL) were cocultured with irradiated allogeneic peripheral blood mononuclear cells (PBMC) in a 7-day MLR. IT2 completely inhibited the proliferative response, comparable to the inhibition with anti-CD28 antibody (Fig. 4B). By contrast, anti-B7 antibody alone only partially inhibited alloantigen-induced MLR, consistent with previous reports^{5,17}. Affinity of the anti-B7 versus anti-B70 antibodies is unlikely to explain the inability of anti-B7 to inhibit primary allogeneic MLR because anti-B7 antibody L307 efficiently blocks MLR when human B7 transfectants are used as stimulators⁷. Therefore, B70 constitutively present on monocytes may initially interact with CD28 on resting T cells to initiate the allogeneic MLR. After the initial stimulation, cytokine production (such as γ -interferon) early during the response may up-regulate B7, which in turn may amplify the response.

We have demonstrated the existence of a second ligand for CD28 and CTLA-4 by generation of an antibody against a unique 70K glycoprotein. The fact that B70 is a ligand for CTLA-4 is demonstrated by the ability of IT2 to inhibit the binding of CTLA4-Ig fusion protein to B-LCL, by the differential immunoprecipitation of B7 and B70 with CTLA4-Ig, and by the direct binding of affinity-purified B70 to CTLA-4. Moreover, the B70 cDNA encodes a novel molecule, distinct from B7, that is able to bind CTLA4-Ig. The functional role of B70 in an immune response is supported by the observation that IT2 completely inhibits primary MLR stimulated with allogeneic PBMC. An intriguing question is why at least two receptors/counter-receptors exist for the CD28/CTLA-4 and B7/B70 costimulation pathway. Redundancy of functional molecules is one explanation, but an alternative possibility is that differential regulation and expression of these genes at various stages of differentiation or activation may be important for the immune system. □

Multi-ion pore behaviour in the CFTR chloride channel

Joseph A. Tabcharani*, Johanna M. Rommens†‡, Yue-Xian Hou‡, Xiu-Bao Chang‡, Lap-Chee Tsui†‡, John R. Riordan‡§ & John W. Hanrahan*||

* Department of Physiology, McGill University, 3655 Drummond Street, Montréal, Québec H3G 1Y6, Canada
Departments of †Molecular and Medical Genetics, and ‡Biochemistry and Clinical Biochemistry, University of Toronto, Toronto, Ontario M5S 1A1, Canada

‡ Research Institute, The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada

Cystic fibrosis transmembrane conductance regulator (CFTR) is a non-rectifying, low-conductance channel^{1,2} regulated by ATP³ and phosphorylation⁴, which mediates apical chloride conductance in secretory epithelia^{5,6} and malfunctions in cystic fibrosis (CF)^{7,8}. Mutations at Lys 335 and Arg 347 in the sixth predicted transmembrane helix of CFTR alter its halide selectivity in whole-cell studies⁹ and its single channel conductance¹⁰, but the physical basis of these alterations is unknown and permeation in CFTR is poorly understood. Here we present evidence that wild-type CFTR can contain more than one anion simultaneously. The conductance of CFTR passes through a minimum when channels are bathed in mixtures of two permeant anions. This anomalous mole fraction effect can be abolished by replacing Arg 347 with an aspartate and can be toggled on or off by varying the pH after the same residue is replaced with a histidine. Thus the CFTR channel should provide a convenient model in which to study multi-ion pore behaviour and conduction. The loss of multiple occupancy may explain how naturally occurring CF mutations at this site cause disease.

Single CFTR Cl⁻ channels were studied in patches excised from Chinese hamster ovary (CHO) cells that had been stably transfected with plasmids directing the expression of wild-type or mutated CFTR. Channels were studied after activation by cyclic AMP-dependent protein kinase A catalytic subunit (PKA; 180 nM) and 1 mM ATP as described previously⁴. A large number of anions were tested for their ability to permeate and for their interfering effects on Cl⁻ current. The traces in Fig. 1a show that adding low concentrations of thiocyanate (SCN⁻) to the cytoplasmic side of wild-type CFTR channels in the presence of 154 mM Cl⁻ decreases single-channel conductance. Figure 1b shows the mean current-voltage relationships obtained with 0, 1, 5 and 10 mM SCN⁻. Thiocyanate on the cytoplasmic side caused a clear outward rectification in the single channel current-voltage (*I/V*) relationship consistent with a fast, voltage-dependent block and also depressed Cl⁻ flow from the 'trans' (extracellular) side. As a first approximation, the electrical distance (δ) of the site and its affinity for SCN⁻ at 0 mV (K_D) were estimated using the Woodhull model for voltage-dependent block¹¹. The results suggest that SCN⁻ ions bind within the channel at a site that is 19.6% of the way through the membrane field from the cytoplasmic side, and that SCN⁻ binding at that site has a dissociation constant of ~ 6.4 mM. There is no reason to assume that this electrical distance corresponds to a physical location. However, an arginine at position 347 in the sixth predicted transmembrane segment (TM6, Fig. 1c) was examined as a potential SCN⁻ blocking site based on previous studies implicating it as a determinant of anion permeation⁹, and because its position (fourth out of 21 amino acids from the cytoplasmic side; 4/21 = 19.1%) is consistent with the electrical distance calculated above. Altering this arginine to a negatively charged aspartate (Arg 347→Asp) abolished block by SCN⁻ (Fig. 1d) and reduced CFTR conductance in symmetrical Cl⁻ solutions to about 53% of that of wild type. Outward

Received 17 September; accepted 15 October 1993.

- Yokochi, T., Holly, R. D. & Clark, E. A. *J. Immun.* **128**, 823-827 (1982).
- Freeman, G. J. et al. *J. Immun.* **143**, 2714-2722 (1989).
- Linsley, P. S., Clark, E. A. & Ledbetter, J. A. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5031-5035 (1990).
- Linsley, P. S. et al. *J. exp. Med.* **173**, 721-730 (1991).
- Koulova, L., Clark, E. A., Shu, G. & Dupont, B. J. *exp. Med.* **173**, 759-762 (1991).
- Azuma, M., Cayabyab, M., Buck, D., Phillips, J. H. & Lanier, L. L. *J. Immun.* **149**, 1115-1123 (1992).
- Azuma, M., Cayabyab, M., Buck, D., Phillips, J. H. & Lanier, L. L. *J. exp. Med.* **175**, 353-360 (1992).
- Domio, N. K., Linsley, P. S. & Ledbetter, J. A. *Eur. J. Immun.* **21**, 1277-1282 (1991).
- Lenschow, D. J. et al. *Science* **257**, 789-792 (1992).
- Linsley, P. S. et al. *Science* **257**, 792-795 (1992).
- Turka, L. A. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 11102-11105 (1992).
- Azuma, M., Cayabyab, M., Phillips, J. H. & Lanier, L. L. *J. Immun.* **150**, 2091-2101 (1993).
- Freedman, A. S., Freeman, G. J., Rhyndhart, K. & Nadler, L. M. *Cell. Immun.* **137**, 429-437 (1991).
- Valle, A., Aubry, J.-P., Durand, I. & Banchereau, J. *Int. Immun.* **3**, 229-236 (1991).
- Azuma, M., Yssel, H., Phillips, J. H., Spits, H. & Lanier, L. L. *J. exp. Med.* **177**, 845-850 (1993).
- Ranheim, E. A. & Kipps, T. J. *J. exp. Med.* **177**, 925-935 (1993).
- Young, J. W. et al. *J. clin. Invest.* **90**, 229-237 (1992).
- Norton, S. D. et al. *J. Immun.* **149**, 1556-1561 (1992).
- Lanier, L. L. & Recktenwald, D. J. *Methods: A Companion to Meth. Enzym.* **2**, 192-199 (1991).
- Aruffo, A. & Seed, B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8573-8577 (1987).
- Lanier, L. L., Rutenberg, J. J. & Phillips, J. H. *J. Immun.* **141**, 3478-3485 (1988).
- Kato, K. et al. *J. exp. Med.* **176**, 1241-1249 (1992).
- Brunet, J.-F. et al. *Nature* **328**, 267-270 (1987).
- Linsley, P. S. et al. *J. exp. Med.* **174**, 561-569 (1991).

ACKNOWLEDGEMENTS. We thank P. Lane, K. Moore, E. Clark, H. Kikutani and H. Spits for generously providing fusion protein, transfectants, antibodies and lymphokines, F. Bazan and D. Nomura for discussion, C. Chang for generation of the cDNA library and discussion, D. Robinson for oligonucleotide synthesis, and D. Polakoff, E. Callias and J. Cupp for assistance with the flow cytometry. DNAX is supported by Schering Plough Corporation. This work was supported by grants from the Ministry of Education and the Ministry of Health in Japan.

|| To whom correspondence should be addressed.

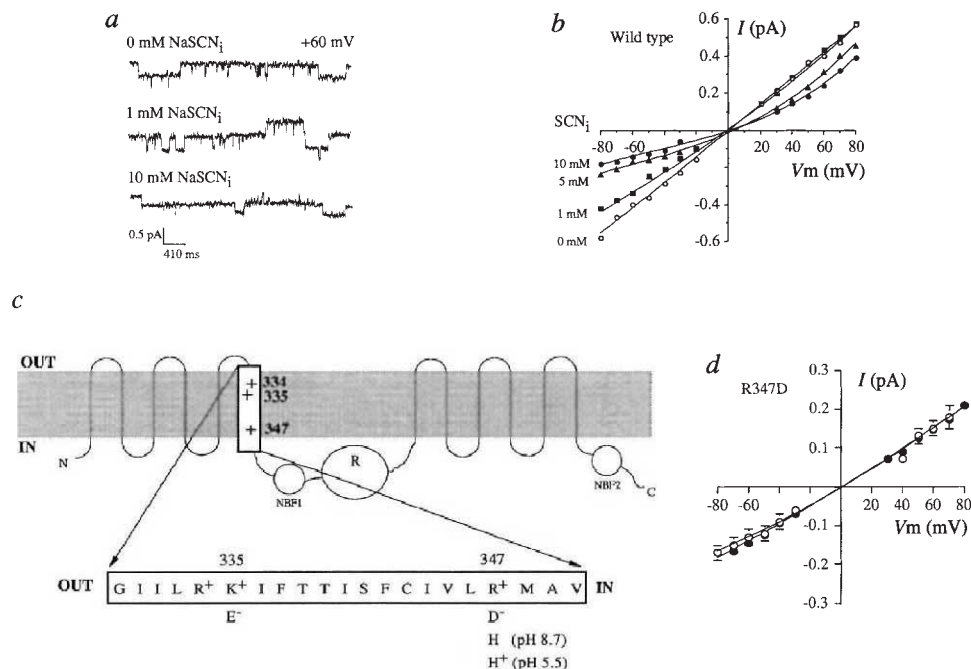
FIG. 1 Effect of thiocyanate on single channel currents. *a*, Traces obtained at +60 mV using membrane patches excised from Chinese hamster ovary cells expressing wild-type CFTR. Channels were activated by exposure to 180 nM protein kinase A catalytic subunit and 1 mM ATP⁸. Patches were bathed with symmetrical 154 mM Cl⁻ solutions under control conditions, or with 1 mM or 10 mM SCN⁻ added to the cytoplasmic side. *b*, Current-voltage relationships determined under control conditions and with increasing concentrations of NaSCN (1–10 mM) on the cytoplasmic side. Lines are least-squares, high-order polynomial fits and are shown for clarity. The electrical distance calculated for the voltage-dependent block using the Woodhull model is 19.6% from the cytoplasmic side, although deviations from the model were observed at elevated SCN⁻ concentrations; rectification was more pronounced and negative currents were somewhat larger than predicted by the theory for a single site. *c*, Diagram showing the relative location and predicted amino-acid sequence for the sixth putative transmembrane spanning region of CFTR. *d*, Current-voltage relationship for the Arg 347→Asp (R347D; single-letter amino-acid code) mutant bathed in control, symmetrical 154 mM Cl⁻ solutions (○) or with 10 mM SCN⁻ added to the cytoplasmic side (●). Note the absence of voltage-dependent SCN⁻ block in the R347D mutant.

METHODS. Plasmids were transfected into CHO-K1 cells and expressing cell lines were selected as described previously^{4,18}. The pNUTCFTF plasmid contains the nucleotide sequences of the entire open reading frame of CFTR under the control of a mouse metallothionein promoter⁴. Site-specific mutagenesis used the polymerase chain reaction (PCR)^{19,20} with Vent polymerase (New England Biolabs, Inc.). Silent changes were initially introduced to create *Stu*I and *Xho*I restriction endonuclease sites at nucleotide positions 950 and 1,279 of the CFTR gene, respectively, to help subsequent mutant construction. The segment containing the mutations made at amino-acid 347 included from *Bsp*EI and *Xho*I restriction sites and for the mutation at amino acid 335 included from *Stu*I to *Fsp*I restriction sites. The manipulated portions of the constructs were verified by sequencing with dideoxy chain termination and Sequenase (US Biochemicals). Immunoreactive CFTR protein was detected in multiple lines generated for each mutant after selection in methotrexate. CFTR was not detected in cells in which plasmid vector alone was integrated. Cells were plated at low density on glass

coverslips and cultured at 37 °C in 5% CO₂ for 1–4 days before use. Experiments were at 20–22 °C. The pipette solution contained (mM) 150 NaCl, 2 CaCl₂, 10 *N*-tris (hydroxymethyl) methyl-2-amino-ethansulphuric acid (TES) (pH 7.4). The bath solution contained (mM) 150 NaCl, 2 MgCl₂, 1 EGTA, 0.5 CaCl₂, 1.0 mM ATP, 180 nM catalytic subunit of cAMP-dependent protein kinase, 10 TES (pH 7.4), and different concentrations of NaSCN as indicated. Type II bovine cardiac PKA was prepared in the laboratory of M. P. Walsh and had a specific activity of 0.45 μmol phosphate incorporated per min per mg enzyme when assayed at 30 °C in 20 mM Tris-HCl (pH 7.5) containing 1 mM EGTA, 5 mM MgCl₂ and 0.2 mM ATP, using histone IIA (Sigma) as substrate. Standard patch-clamp techniques were used to generate data in this and subsequent figures. PKA-stimulated channels were not observed in untransfected CHO cells, and only low-conductance Cl⁻ channels were consistently activated by PKA in transfected cells. Recordings were low-pass filtered at 150 Hz (8-pole Bessel) and digitized at 1 kHz. Current amplitudes were measured using a semi-automated procedure that involved computing amplitude histograms for short segments of record and displaying them sideways on a split screen next to the raw data so that peaks could be verified using cursors. Event amplitudes were measured at each potential, and displayed as an *I/V* curve at the end of the run. Other details of data acquisition and analysis were described previously^{6,21}.

rectification was surprisingly weak in the Arg 347→Asp mutant considering the drastic alteration in charge. This implied that low conductance in Arg 347→Asp might be caused by some fundamental change in the permeation pathway rather than simple addition of a new barrier to anions (such as Asp⁻). Channels can hold more than one ion simultaneously (reviewed in ref. 12), and the destabilization which results from ion-ion repulsion inside multiply occupied pores is thought to enable ions to pass through rapidly despite high-affinity binding^{13,14}. Strong evidence for a multi-ion pore can be obtained by measuring changes in conductance and permeability ratio in solutions containing two permeant ions; conductance of a single ion pore should increase monotonically as the mole fraction of the more permeant species is increased. In a multi-ion pore, conductance may decrease in mixtures when compared to pure solutions containing either ion.

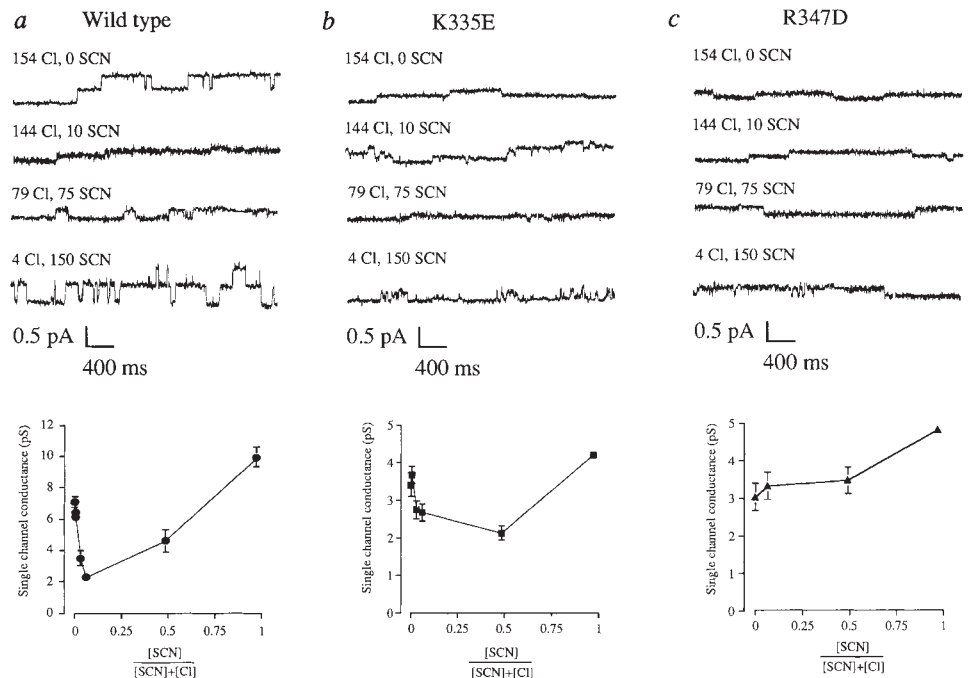
Figure 2*a* shows that wild-type CFTR exhibits a classic anomalous mole fraction effect when bathed symmetrically in mixtures of Cl⁻ and SCN⁻. Conductance decreased from 6.9 pS at room temperature in symmetrical 154 mM Cl⁻ solutions to 2.2 pS



when 10 mM Cl⁻ was replaced by SCN⁻, and then increased again to 10.3 pS at 150 mM SCN⁻. This anomalous drop indicates there are interactions between permeating anions and therefore implies that at least two anions can reside simultaneously inside the pore.

We examined the importance of two positively charged residues in TM6 for anomalous mole fraction effects by mutating them to acidic amino acids. When Lys 335, another residue that has been proposed to interact with permeating anions⁹, was changed to Glu, multi-ion pore behaviour was still observed (Fig. 2*b*), although the concave relationship between conductance and mole fraction was more shallow and shifted to somewhat higher SCN⁻ concentrations. By contrast, altering Arg 347 to an aspartate abolished anomalous mole fraction effects (Fig. 2*c*). Single-channel conductance in symmetrical Cl⁻ solutions was reduced compared to wild type, and it increased monotonically as the thiocyanate mole fraction was increased. The absence of anomalous mole fraction effects in Arg 347→Asp suggests that ion-ion interactions within the pore are lost after this mutation. The simplest explanation for this change in behav-

FIG. 2 Anomalous mole fraction effects in wild-type and mutated CFTR channels. **a**, Wild-type, traces were obtained at +50 mV with different mixtures of chloride and thiocyanate and the relationship between single-channel conductance and thiocyanate mole fraction $[\text{SCN}^-]/([\text{SCN}^-]+[\text{Cl}^-])$ in symmetrical solutions is shown below. Chord conductance at +50 mV passed through a sharp minimum and then increased as chloride was progressively substituted by thiocyanate. The highest SCN^- mole fraction tested was 0.974 (150 mM SCN^- : 4 mM Cl^-). Mole fraction effects on the permeability ratio were not characterized because transitions were not observed at negative potentials after exposure to most SCN^-/Cl^- mixtures on either side and because currents near the reversal potentials were too small to resolve (<0.1 pA), necessitating large extrapolations (30 mV). **b**, Lys 335→Glu (K335E), same conditions as in **a**, except patches were excised from CHO cells expressing the K335E mutant. Single-channel conductance was reduced to less than 4 pS in pure Cl^- solutions and the current-voltage relationship, which was normally linear in excised patches, became slightly rectified in the inward direction as expected versus wild type shown in Fig. 1d). **c**, R347D: using identical conditions as for **a** and **b**,



ior would be the loss of multiple occupancy, although the present data do not exclude other mechanisms.

To examine specifically the role of positive charge at amino-acid position 347, we mutated the arginine to a histidine (Arg 347→His). Histidine would be predominantly positively charged at low pH but uncharged at high pH. Figure 3a shows traces obtained when the mutant was bathed with symmetrical 154 mM Cl^- solutions at pH 5.5. Under these conditions, the single-channel conductance (6.34 pS) and voltage-dependent block by 10 mM thiocyanate on the cytoplasmic side (Fig. 3b) were identical to those observed with wild-type CFTR. The conductance ratio at +50 mV measured in the presence versus absence of 10 mM SCN^- was 0.34 ± 0.03 (mean $\pm$ s.e.; $n=5$) for wild-type CFTR and 0.32 ± 0.02 ($n=4$) for Arg 347→His at pH 5.5. When the pH on both sides of the membrane was increased to 8.7 so that the histidine would be predominantly in its neutral form, single-channel current amplitude was reduced by $\sim 60\%$ (Fig. 3c) and the channel was no longer blocked by SCN^- (Fig. 3d). Closely similar results were obtained with Arg 347→His at high pH and with Arg 347→Asp at pH 7.4; therefore the low conductance and absence of SCN^- block in Arg 347→Asp were not caused by aspartate's smaller size or its negative charge, but rather was due to the lack of positive charge, consistent with Arg 347 being a site (free energy well) that binds permeating anions by electrostatic interactions. Considering the importance of Arg 347 for the mole fraction effects on conductance (Fig. 2), we examined whether anomalous mole fraction effects would be pH-dependent in Arg 347→His. Single-channel conductance of Arg 347→His was measured at pH 5.5 and 8.7 using the $\text{Cl}^-:\text{SCN}^-$ mixture that yielded the strongest anomalous mole fraction effect in wild-type CFTR (0.065 SCN^- ; Fig. 2a). Figure 4 shows the anomalous mole-fraction effect was indeed pH-dependent in the mutant. The conductance of wild-type CFTR and Arg 347→Asp in this anion mixture at pH 7.4 are also shown for comparison. At low pH, where the His 347 in the mutant would be positively charged, conductance was reduced by about threefold in $\text{Cl}^-:\text{SCN}^-$ mixtures, a decrease that was nearly identical to the one observed for wild-type CFTR under

conductance was significantly reduced in symmetrical Cl^- solutions compared to the K335E mutant and increased monotonically with SCN^- mole fraction.

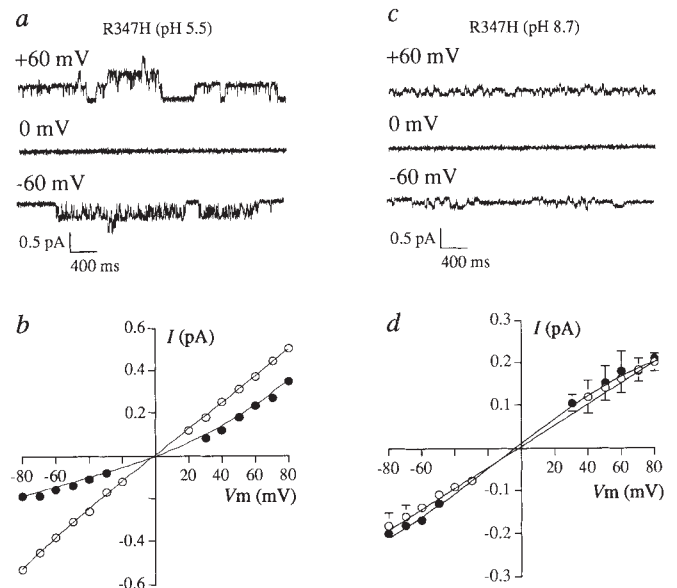


FIG. 3 Effect of pH on voltage-dependent thiocyanate block when Arg 347 was replaced by His (R347H). **a**, Single-channel traces of R347H obtained at pH 5.5 in the absence of thiocyanate. **b**, I/V relationship for R347H at pH 5.5 in symmetrical 154 Cl^- solutions ($\circ$) and with 10 mM SCN^- added to the cytoplasmic side ($\bullet$). **c**, Traces of R347H obtained at pH 8.7 in the absence of SCN^- . **d**, I/V relationship observed in the absence ($\circ$) or presence ($\bullet$) of 10 mM SCN^- on the cytoplasmic side.

comparable conditions. In contrast, when Arg 347→His was exposed to the same SCN^- mole fraction at pH 8.7, conductance increased slightly as expected from the higher conductance and permeability ratios for SCN^- in CFTR ($P_{\text{SCN}^-}/P_{\text{Cl}^-} = 1.5 \pm 0.11$).

These results suggest that anions interact within the CFTR pore and therefore imply that the channel can be simultaneously occupied by more than one anion. One of the SCN^- and pre-

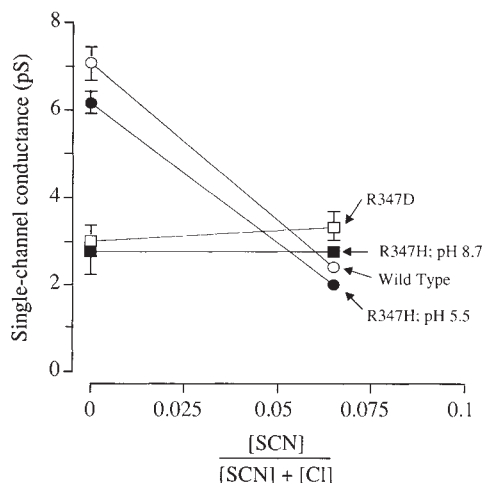


FIG. 4 Anomalous mole fraction effect in the R347H mutant toggled by pH. At pH 5.5, the symmetrical $\text{SCN}^-:\text{Cl}^-$ mixture caused a large decrease in conductance (●), similar to that observed with wild-type CFTR channels studied at pH 7.4 (○). By contrast, the conductance of R347H was reduced by more than 50% in pure Cl^- solutions when pH was increased to 8.7, and conductance did not decrease in the $\text{Cl}^-:\text{SCN}^-$ mixture (■). These results, obtained when the histidine would be predominantly uncharged, resemble those obtained with R347D at neutral pH (□, and Fig. 2), when the site would have a negative charge.

sumably Cl^- sites may be near Arg 347. Flux ratio experiments are needed to establish single occupancy in mutants lacking a positive charge at position 347. Nevertheless, the present results are consistent with model predictions for other pores^{13,14} that multiple occupancy increases conductance, in that mutations at amino-acid position 347 that eliminate the anomalous mole fraction dependence of conductance also reduce the conductance by about 50%.

Arg 347 is the site of several naturally occurring mutations (R347P¹⁵, R347L¹⁶, R347H¹⁷; Arg 347→Pro, Leu and His), which are associated with relatively mild cystic fibrosis¹⁵. Moreover, it has recently been reported that the conductance of Arg 347→Pro is reduced from 7.9 pS to 2.34, remarkably similar to the values of 2.7 pS (Arg⁺ 347→Asp⁻) and 2.43 pS (Arg⁺ 347→His, pH 8.7) obtained here. Based on the present findings, we speculate that mutations at Arg 347 reduce conductance and cause cystic fibrosis by converting CFTR from a multi-ion to a single-ion pore. □

Received 14 June; accepted 6 August 1993.

- Kartner, N. et al. *Cell* **64**, 681–691 (1991).
- Bear, C. E. et al. *Cell* **68**, 809–818 (1992).
- Anderson, M. P. et al. *Cell* **67**, 775–784 (1991).
- Tabcharani, J. A., Chang, X.-B., Riordan, J. R. & Hanrahan, J. W. *Nature* **352**, 628–631 (1991).
- Gray, M. A., Harris, A., Coleman, L., Greenwell, J. R. & Argent, B. E. *Am. J. Physiol.* **257**, C240–C251 (1989).
- Tabcharani, J. A., Low, W., Elie, D. & Hanrahan, J. W. *FEBS Lett.* **270**, 157–164 (1990).
- Rich, D. P. et al. *Nature* **347**, 358–363 (1990).
- Drumm, M. L. et al. *Cell* **62**, 1227–1233 (1990).
- Anderson, M. P. et al. *Science* **253**, 202–205 (1991).
- Sheppard, D. N. et al. *Nature* **362**, 160–164 (1993).
- Woodhull, A. M. *J. gen. Physiol.* **81**, 687–708 (1973).
- Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, Mass., 1992).
- Almers, W. & McCleskey, E. W. *J. Physiol. Lond.* **353**, 585–608 (1984).
- Hess, P. & Tsien, R. W. *Nature* **309**, 453–456 (1984).
- Kristidis, P. et al. *Am. J. hum. Genet.* **50**, 1178–1184 (1992).
- Audrézet, M. P. et al. *Hum. molec. Genet.* **2**, 51–54 (1993).
- Cremonesi, L. et al. *Hum. Mut.* **1**, 314–319 (1992).
- Chang, X.-B. et al. *J. Biol. Chem.* **268**, 11304–11311 (1993).
- Higuchi, R., Krummel, B. & Saiki, R. K. *Nucleic Acids Res.* **16**, 7351–7363 (1988).
- Ho, S. N., Hunt, H. D. & Horton, R. M. *Gene* **77**, 51–59 (1989).
- Tabcharani, J. A., Jensen, T. J., Riordan, J. R. & Hanrahan, J. W. *J. membr. Biol.* **112**, 109–122 (1989).

ACKNOWLEDGEMENTS. We thank V. Lai and A. Boucher for technical assistance. This work was supported by the Canadian and US CF Foundations, the NIH (NIDDK) and the Respiratory Health Network of Centres of Excellence (CF Theme). J.W.H. is a Chercheur-Boursier of the Fonds de la recherche en santé du Québec.

The mitotic feedback control gene *MAD2* encodes the α -subunit of a prenyltransferase

Rong Li^{*†}, Christopher Havel^{*}, John A. Watson^{*} & Andrew W. Murray^{*‡§}

Departments of ^{*} Biochemistry and [†] Physiology, University of California, San Francisco, California 94143-0444, USA

THE *mad2-1* mutation inactivates the cell-cycle feedback control that prevents budding yeast cells from leaving mitosis until spindle assembly is complete¹. The gene product of *MAD2* shows significant sequence similarity to the α -subunit of prenyltransferases². Here we isolate a new temperature-sensitive *mad2* mutant, *mad2-2^{ts}*, and find that Mad2p is required for the membrane association of Ypt1p and Sec4p, two prenylated small GTP-binding proteins involved in protein trafficking. Extracts from *mad2-2^{ts}* mutant cells fail to geranylgeranilate a number of substrates at the non-permissive temperature. *mad2-2^{ts}* is synthetically lethal with *bet2-1*, a mutation in the gene that encodes for the β -subunit of the Ypt1p and Sec4p geranylgeranyl transferase^{3,4}. Therefore *MAD2* and *BET2* gene products may physically interact to form a geranylgeranyl transferase complex. In addition, the difference between the phenotypes of *mad2-1* and *mad2-2^{ts}* suggests that *MAD2* has distinct roles in protein transport and the mitotic feedback control.

To investigate whether Mad2p plays a role in prenylating cellular proteins, we exploited the observation that the membrane localization of many small GTP-binding proteins requires prenylation^{5–7}. To detect GTP-binding proteins, extracts from wild-type and *mad2-2^{ts}* cells were separated into membrane-bound and soluble fractions, separated on SDS-polyacrylamide gels, transferred onto nitrocellulose and incubated with [α -³²P]GTP⁸. Two prominent proteins of *M_r* 23–24K which were primarily in the membrane fraction in wild-type cells and *mad2-2^{ts}* cells grown at 23 °C, shifted to the soluble fraction in *mad2-2^{ts}* cells grown at 37 °C (Fig. 1a). This change in subcellular localization suggests that Mad2p may be required for the prenylation of these two small GTP-binding proteins. In contrast, extracts prepared from *mad2-1* cells gave the same distribution of GTP-binding proteins as wild-type cells, suggesting that the feedback control defect of this mutant is probably not due to a change in the localization of any of the GTP-binding proteins detected by this relatively crude assay.

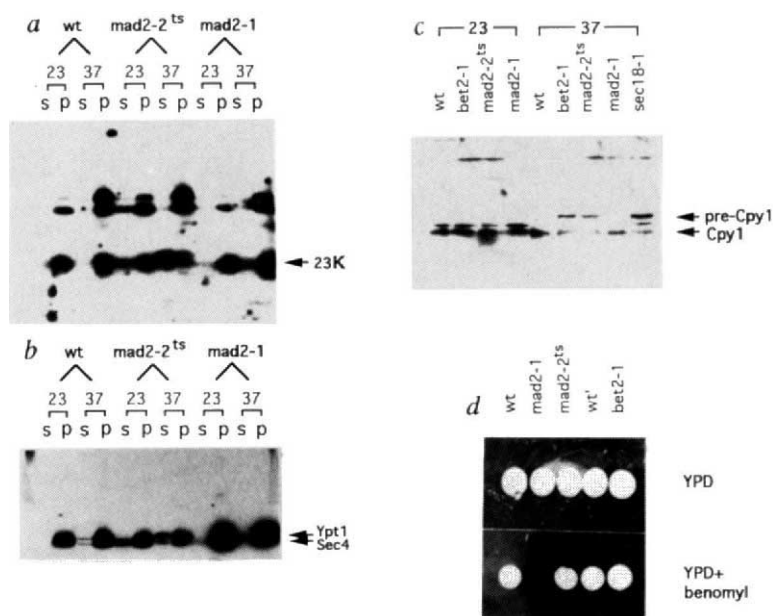
The relative molecular masses of the GTP-binding proteins that appeared in the soluble fraction in *mad2-2^{ts}* cells at 37 °C are the same as those of Ypt1p and Sec4p, two small GTP-binding proteins involved in protein trafficking^{9–12}. Immunoblots with antibodies recognizing Ypt1p or Sec4p demonstrated that a fraction of these proteins become soluble in *mad2-2^{ts}* strains at both 23 and 37 °C (Fig. 1b). GTP binding experiments showed that most of the activity in this size region becomes soluble. The discrepancy between the GTP blot and immunoblot results could be due to a number of factors: the presence of other small G proteins in addition to Ypt1p and Sec4p in the 23–24K range that account for most of the GTP binding, or preferential GTP binding to, or renaturation of, unprenylated G proteins.

Because the membrane attachment and wild-type functions of Ypt1p and Sec4p depend on Bet2p, a prenyltransferase β -subunit, *mad2-2^{ts}* mutant cells should have a defect in protein trafficking similar to that seen in *bet2-1* if Mad2p is also required for Ypt1p and Sec4p prenylation. We examined the effect of *mad2-2^{ts}* on carboxypeptidase Y (Cpy1p) maturation, because its

[†] Present address: Molecular and Cellular Biology, University of California, Berkeley, California 94720, USA.

[§] To whom correspondence should be addressed.

FIG. 1 *mad2-2^{ts}* mutants fail to localize small GTP-binding proteins to membrane-bound compartments. **a**, Extracts prepared from wild-type, *mad2-2^{ts}* and *mad2-1* cells incubated at either 23 or 37 °C were separated into soluble (s) and membrane (p) fractions. Proteins from each fraction were resolved on an SDS-polyacrylamide gel, transferred to nitrocellulose, and incubated with [α -³²P]GTP to detect GTP-binding proteins. **b**, The same blot was immunoblotted sequentially with antibodies recognizing Ypt1p and Sec4p. **c**, Western analysis with antibody to Cpy1p was done on extracts prepared from wild-type, *bet2-1*, *mad2-2^{ts}* and *mad2-1* cells grown at 23 or 37 °C for 3 h. **d**, Equal amounts of *mad2-2^{ts}*, *mad2-1*, their congenic wild-type (wt), *bet2-1* and its congenic wild-type (wt') cells were spotted on YPD or YPD-containing 15 μ g ml⁻¹ benomyl plates. The plates were photographed after a 4-day incubation at 23 °C. **METHODS**. Exponentially growing cultures were collected after growth at 23 °C, or after incubation for 3 h at 37 °C. Spheroplast formation and cell fractionation were as described¹⁹. Equal amounts of proteins from each fraction were run on the gel. GTP-binding assays were done as described⁸. Extracts for the western analysis using Cpy1p antibody were prepared by glass-bead lysis of 1.0 OD₆₀₀ units of cells in standard SDS sample buffer followed by boiling for 5 min. To obtain a temperature-sensitive *mad2* allele, we mutagenized plasmid pRL19, which carries the *MAD2* gene, using hydroxylamine²⁰, and screened for plasmids that allowed $\Delta*mad2* cells to grow at room temperature but not at 37 °C.$



Plasmid shuffling²⁰ was then done to construct the strain that carries the chromosomal *mad2-2^{ts}* allele.

processing and vacuolar location depends on *YPT1*¹³. In wild-type or *mad2-1* cells, most of Cpy1p is in its mature, vacuolar form, but *mad2-2^{ts}* and *bet2-1* cells grown at 37 °C accumulate much of the unprocessed precursor (Fig. 1c).

The lack of Ypt1p and Sec4p membrane attachment and the similar defect in Cpy1p transport in *mad2-2^{ts}* and *bet2-1* cells suggest that Mad2p and Bet2p may interact to form an active prenyltransferase. We mated a *bet2-1* strain to a *mad2-2^{ts}* strain, sporulated the resulting diploid and did tetrad analysis. No *mad2-2^{ts}* *bet2-1* double mutant spores were recovered (Table 1), demonstrating that the two mutations are synthetically lethal and supporting the idea that the gene products interact with each other. Strong biochemical support for this hypothesis is provided in the accompanying letter¹⁴.

TABLE 1 *bet2-1* is synthetically lethal with *mad2-2^{ts}*

	4-Spore tetrads		3-Spore tetrads		2-Spore tetrads					
ts: wt →	4:0	3:1	2:2	3:0	2:1	1:2	2:0	1:1	0:2	
Number of tetrads	12	1	0	3	49	0	1	0	25	

A diploid strain obtained by crossing *bet2-1* and *mad2-2^{ts}* strains was sporulated. 100 tetrads were dissected, and the segregation of the temperature-sensitive (ts) phenotype was scored.

TABLE 2 Relative amounts of farnesylated peptides and geranylgeranylated peptides in wt and *mad2-2^{ts}* extracts

	Farnesylated peptides	Geranylgeranylated peptides
wt	11%	89%
<i>mad2-2^{ts}</i>	81%	19%

Extracts prepared from wild-type (wt) and *mad2-2^{ts}* cells as described in Fig. 2 legend were pre-incubated at 37 °C for 30 min. Prenylation reactions were done at 37 °C and were then delipidized and trypsinized¹⁶. The resulting prenylated peptides were analysed by radio-labelled reverse-phase HPLC¹⁶ with a 50–100% ACN gradient for 30 min followed by an isocratic for 10 min. The percentages of farnesylated peptides (retention time, 10–20 min) and geranylgeranylated peptides (retention time, 20–31 min) are shown.

To determine the substrates of the Mad2p-containing prenyltransferase, we developed a method of labelling endogenous yeast prenylated proteins. Cells that accumulated unprenylated protein substrates were obtained by treatment with lovastatin, which inhibits the production of mevalonic acid, the precursor of farnesyl and geranylgeranyl groups¹⁵. Extracts from these cells can use exogenously added ³H-mevalonic acid to synthesize labelled farnesyl and geranylgeranyl pyrophosphates and prenylate existing proteins. The labelled proteins were resolved on specialized two-dimensional polyacrylamide gels that had been used effectively to resolve prenylated proteins¹⁶. The prenylation products in wild-type extracts resolved into more than ten distinct spots, many of which co-migrate with GTP-binding activities (Fig. 2a, b, g). In *bet2-1* cells, the pattern of labelling at room temperature resembled that of wild-type cells, but extracts pre-incubated at 37 °C failed to label several proteins (Fig. 2e, f). The labelling defects in *mad2-2^{ts}* were more severe: even at room temperature the labelling of many low-*M_r* proteins was sharply reduced and pre-incubation of the extracts at 37 °C before the labelling reaction completely abolished the labelling of these spots (Fig. 2c, d).

To determine which prenyl group is transferred by the Mad2p-containing prenyltransferase, the prenylated proteins in wild-type and *mad2-2^{ts}* extracts were proteolytically digested and the peptides were separated on reverse-phase high performance liquid chromatography (HPLC). In wild-type cells most of the radioactivity was recovered in geranylgeranylated rather than farnesylated peptides. In *mad2-2^{ts}* extracts that had been pre-incubated at 37 °C, however, most of the radioactivity is in farnesylated peptides (Table 2). This observation, coupled with the demonstration that the synthesis of geranylgeranyl pyrophosphate is unaffected in *mad2-2^{ts}* extracts (data not shown), suggests that Mad2p is required for most geranylgeranyltransferase activity in yeast. The α - and β -subunits of the mammalian type II geranylgeranyltransferase are most closely related to Mad2p and Bet2p, suggesting that these enzymes may have similar functions in mammalian cells and yeast¹⁷.

We have not yet identified a completely reproducible difference in the pattern of prenylation in extracts prepared from wild-type and *mad2-1* cells. Because *mad2-1* cells are not defective in Ypt1p and Sec4p localization, and neither *bet2-1* nor *mad2-2^{ts}* has defects in the mitotic feedback control, as monitored by

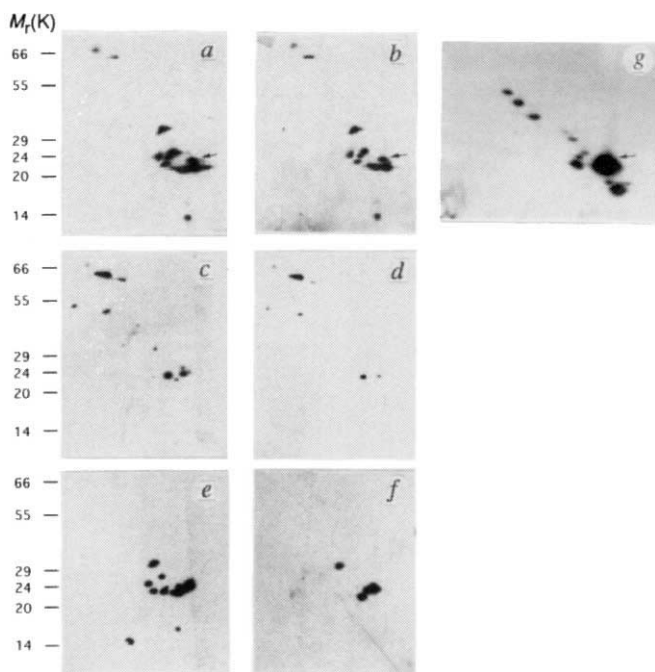


FIG. 2 A novel assay for prenyltransferase activities in yeast extracts revealed that Mad2p is a component of a geranylgeranyltransferase. a–f, The pattern of labelling in wild-type (a, b), *mad2-2^{ts}* (c, d) and *bet2-1* (e, f) cells. g, Two-dimensional GTP blot. The labelling was done under two conditions: 30-min preincubation at 23 °C with labelling reaction at 23 °C (a, c, e), or 30-min pre-incubation at 37 °C with labelling at 37 °C (b, d, f). Arrows in a, b and g indicate the position of Ypt1p and Sec4p as determined by immunoblots (data not shown).

METHODS. Lovastatin (0.1 mg ml⁻¹; a gift from Merck Research Laboratories) was added to 100 ml of exponentially growing cultures, OD₆₀₀ = 0.3. After 3 h at 23 °C, spheroplasts were prepared and lysed in 200 µl labelling buffer (50 mM potassium phosphate, pH 7.5, 50 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT and 1 mM PMSF) as described¹⁹. The lysate was spun for 15 min at 15,000g and the supernatant was used for *in vitro* prenylation. 10 µl of 6 × reaction buffer was added to 50 µl of extracts to yield the following final concentrations: 5 mM MgCl₂, 0.02 mg ml⁻¹ ZnSO₄, 5 mM ATP, 12.5 mM phosphoenolpyruvate (Sigma), 20 units of pyruvate kinase (Sigma) and 250 µCi [³H]-mevalonate (12.7 mCi µmol⁻¹), synthesized as described²¹. Two-dimensional polyacrylamide electrophoresis with BAC (benzylidimethyl-*n*-hexadecylammonium chloride; Sigma) gels as the first dimension and SDS gels as the second dimension was as described¹⁶. The same unlabelled wild-type extract was electrophoresed, transferred to nitrocellulose, and incubated with [³²P]GTP as described⁸, except that 0.1% BSA and 30 µM ATP were added to the GTP binding buffer to reduce the background.

benomyl-sensitivity at 23 °C (Fig. 1d), we suggest that Mad2p, like other prenyl transferase α -subunits^{4,18}, forms complexes with more than one β -subunit. The *mad2-2^{ts}* mutation would affect Mad2p-Bet2p complexes responsible for the prenylation of Ypt1p and Sec4p, whereas the *mad2-1* mutation would affect complexes between Mad2p and another β -subunit responsible for prenylating an essential component of the mitotic feedback control. The observations that the *in vitro* labelling of some proteins is dependent on *MAD2* but not on *BET2* (Fig. 2) supports the hypothesis that Mad2p can associate with more than one β -subunit. Nevertheless, we cannot exclude the possibility that the role of Mad2p in the mitotic feedback control is unrelated to its prenyltransferase activity. □

Received 23 April; accepted 11 August 1993.

- Li, R. & Murray, A. W. *Cell* **66**, 519–531 (1991).
- Boguski, M., Murray, A. W. & Powers, S. *New Biol.* **4**, 408–411 (1992).
- Rossi, G., Jiang, Y., Newman, A. P. & Ferro-Novick, S. *Nature* **351**, 158–161 (1991).
- Kohl, N. E. et al. *J. Biol. Chem.* **266**, 18884–18888 (1991).

- Finegold, A. A., Schafer, W. R., Rine, J., Whiteway, M. & Tamanoi, F. *Science* **249**, 165–169 (1990).
- Schafer, W. R. et al. *Science* **249**, 1133–1139 (1990).
- Glomset, J., Gelb, M. & Farnsworth, C. *Curr. Opin. Lipidol.* **2**, 118–124 (1992).
- Barnes, G., Louie, K. A. & Botstein, D. *Molec. Biol. Cell* **3**, 29–47 (1992).
- Bacon, R. A., Salminen, A., Ruohola, H., Novick, P. & Ferro, N. S. *J. Cell Biol.* **109**, 1015–1022 (1989).
- Segev, N. *Science* **252**, 1553–1556 (1991).
- Walworth, N. C., Goud, B., Kabacoff, A. K. & Novick, P. *J. EMBO J.* **8**, 1685–1693 (1989).
- Moya, M., Roberts, D. & Novick, P. *Nature* **361**, 460–463 (1993).
- Newman, A. & Ferro-Novick, S. *J. Cell Biol.* **105**, 1587–1594 (1987).
- Jiang, Y., Rossi, G. & Ferro-Novick, S. *Nature* **366**, 84–86 (1993).
- Alberts, A. W. et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3957–3961 (1980).
- Havel, C. M. & Watson, J. A. *Arch. Biochem. Biophys.* **295**, 410–420 (1992).
- Armstrong, S. A., Seabra, M. C., Sudhof, T. C., Goldstein, J. L. & Brown, M. S. *J. Biol. Chem.* (in the press).
- Seabra, M. C., Reiss, Y., Casey, P. J., Brown, M. S. & Goldstein, J. L. *Cell* **65**, 429–434 (1991).
- Deshaies, R. J. & Schekman, R. **109**, 2653–2664 (1989).
- Guthrie, C. & Fink, G. R. *Meth. Enzym.* **194**, 315–318 (1991).
- Keller, R. K. *J. Biol. Chem.* **261**, 12053–12059 (1986).

ACKNOWLEDGEMENTS. We thank Y. Jiang, G. Rossi and S. Ferro-Novick for advice, helpful discussions, communicating unpublished results, *bet-2* strains and antibodies to Ypt1p, P. Novick for antibodies to Sec4p, staff from the Schekman laboratory for antibody to Cpy1p, J. Goldstein for communicating unpublished results, C. Trueblood, R. Deshaies, T. Mitchison, E. Blackburn and members of our laboratories for stimulating discussions, and D. Drubin, G. Barnes, D. Sun, J. Minshull and R. Vale for comments and help during manuscript preparation. A. W. M. is a Lucille P. Markey Scholar and David and Lucile Packard fellow. This work was supported by grants from the NIH, the Markey Charitable Trust, the Packard foundation and the March of Dimes.

Bet2p and Mad2p are components of a prenyltransferase that adds geranylgeranyl onto Ypt1p and Sec4p

Yu Jiang, Guendalina Rossi & Susan Ferro-Novick*

Department of Cell Biology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA

THREE different prenyltransferases have been identified in yeast and higher cells^{1–6}, the farnesyltransferase and the type I and type II geranylgeranyltransferases (GGTase). The farnesyltransferase and GGTase-I modify peptides *in vitro* with the CAAX (C, Cys; A, aliphatic residue; X, terminal amino acid) consensus motif^{2,7}. These enzymes are heterodimers that have different β -subunits and a shared α -subunit⁸. In yeast, the *RAM2* gene encodes this α -subunit⁹. *RAM2* is also homologous to *MAD2*, a yeast gene whose product has been implicated in the feedback control of mitosis^{10,11}. We have shown that Bet2p is a component of the yeast GGTase-II (refs 6, 12) that geranylgeranylates Ypt1p, a small GTP-binding protein that mediates transport from the endoplasmic reticulum to the Golgi complex^{13–15}. Here we report that Mad2p is a component of this enzyme. Bet2p forms a complex with Mad2p that appears to bind geranylgeranyl pyrophosphate, but not farnesyl pyrophosphate. The efficient transfer of geranylgeranyl onto small GTP-binding proteins requires the presence of an additional activity.

We used an *in vitro* prenylation assay to determine whether Sec4p, a GTP-binding protein that functions in post-Golgi secretion¹⁶, is a substrate for the yeast GGTase-II. Under conditions in which Ypt1p was geranylgeranylated *in vitro* (Fig. 1, lane 1), Sec4p was also geranylgeranylated (Fig. 1, lane 5). Furthermore, a *bet2* mutant extract (100,000g supernatant) was defective in this activity (Fig. 1, lanes 2 and 6; see quantification in Fig. 2a, b). Ypt1p and Sec4p were not prenylated when the

* To whom correspondence should be addressed.

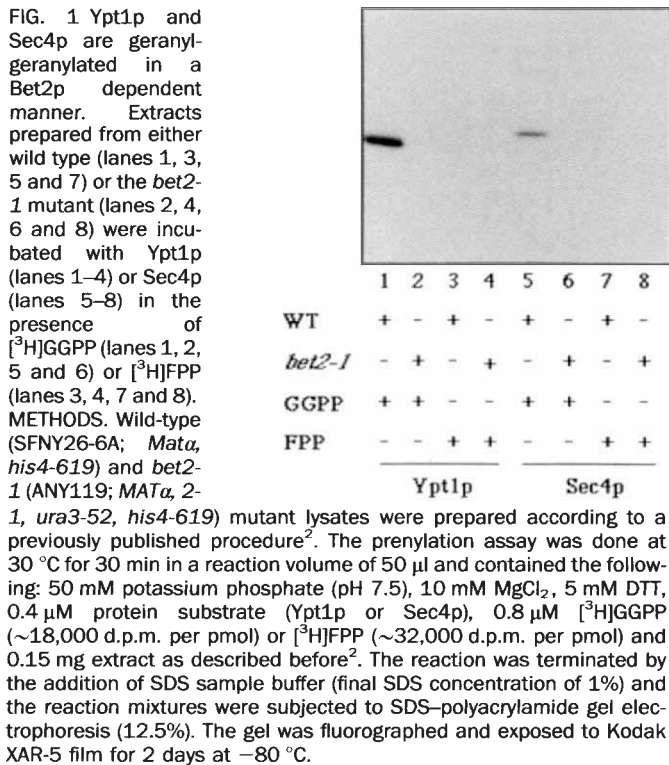


FIG. 2 Geranylgeranylation of Sec4p and mutant forms of Ypt1p. Geranylgeranylation activity of wild-type and *bet2-1* mutant lysates was determined using *E. coli*-expressed Ypt1p, Sec4p and mutant forms of Ypt1p.

METHODS. Mutant forms of the *YPT1* gene were constructed by replacing the codons at the C terminus (Cys-Cys) with either Cys-Ser-Cys (Ypt1p-CXC) or Cys-Thr-Ile-Leu (Ypt1p-CAAL) using oligonucleotide-directed mutagenesis. These mutant *YPT1* genes were cloned into a T7 expression plasmid (pET-11d) by polymerase chain reaction (PCR) and expressed in BL21(DE3) bacterial cells. The protocol that was used for the expression and purification of Ypt1p and Sec4p has already been described^{26,27}.

Geranylgeranyl:protein transferase activity is indicated as the amount of [³H]geranylgeranyl that is transferred from [³H]GGPP into acid-precipitable material. The assay was as described in the legend to Fig. 1. The reaction was terminated by the addition of 1 M HCl in ethanol (1 ml) and then filtered on a Whatman GF/A filter. The filter was repeatedly washed with ethanol, dried, and counted in a scintillation counter. The blank (reaction mix minus substrate) was ~1,000 cpm and was subtracted from each sample. The units of activity are expressed as pmol of [³H]geranylgeranyl that is attached to Ypt1p, Sec4p, and mutant forms of Ypt1p per min per mg of extract. Wild type (open squares); *bet2-1* (black diamonds).

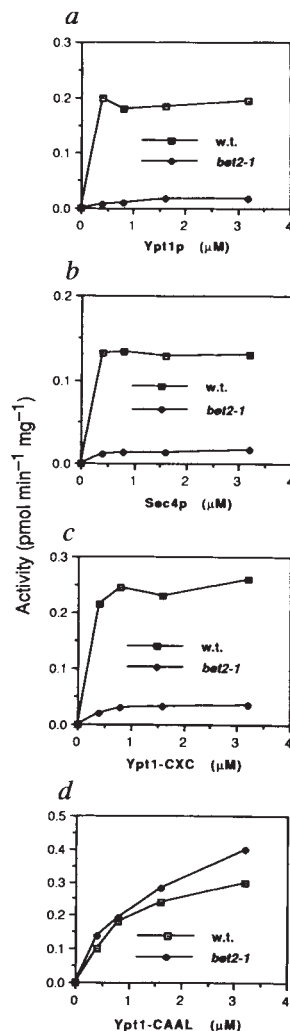


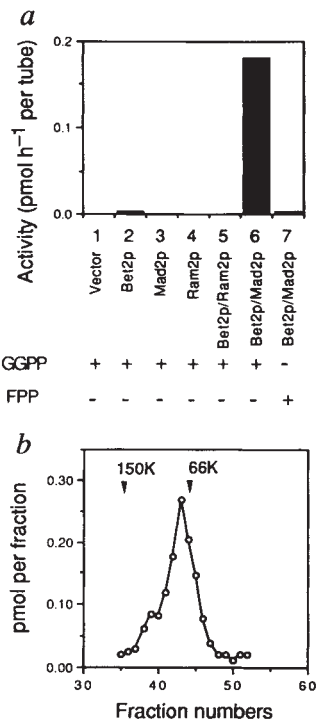
FIG. 3 Co-expressed Bet2p/Mad2p form a 70K complex. a, The activity of extracts prepared from cells expressing different constructs was determined. Either [³H]GGPP (lanes 1–6), or [³H]FPP (lane 7), was used as the lipid donor. Lane 1, vector without insert; lane 2, Bet2p ne; lane 3, 2p alone; lane 4, Ram2p alone; lane 5, Bet2p and Ram2p co-expressed together; lanes 6 and 7, co-expressed Bet2p and Mad2p. b, Gel filtration chromatography indicated that the activity migrated with an apparent *M_r* of 70K.

METHODS. To express *BET2* in *E. coli*, the coding sequence of the *BET2* gene was generated by PCR using two primers that overlapped with the initiation or stop codon. A *KpnI* restriction enzyme site was incorporated into the 5' primer and a *Sall* site was introduced into the 3' primer. This DNA fragment was digested with *KpnI* and *Sall* and then cloned into an ampicillin-resistant plasmid (pUC118) that was cut with the same enzymes. The *RAM2* gene in plasmid pBC-KS (pBH57) was a gift from S. Powers⁹. For the expression of *MAD2* in *E. coli*, a PCR product that contained the *MAD2* coding sequence was digested with *KpnI* and *ClaI* and cloned into a chloramphenicol-resistant plasmid (pBC-KS) that was cleaved with the same enzymes. A 0.439-kb *KpnI*-*AflIII* fragment was removed from this vector and replaced with a 0.336-kb *KpnI*-*AflIII* fragment that was obtained from pUC118. This operation reduced the region that resided upstream of *MAD2*. Constructs containing *BET2* and *MAD2* (or *RAM2*) were simultaneously transformed into JM101 bacterial cells and expressed as described before⁹.

Activity was determined by measuring the amount of [³H]GGPP that was trapped onto a Whatman filter. The reaction was done in a final volume of 50 μl and contained the following: 50 mM Tris-chloride (pH 7.5), 20 μM ZnCl₂, 10 mM MgCl₂, 5 mM DTT, 0.4 μM Ypt1p, 0.8 μM [³H]GGPP (~18,000 d.p.m. per pmol), or [³H]FPP (~32,000 d.p.m. per pmol), and 50 μg extract. After a 30 min incubation at 30 °C, the reaction was terminated by the addition of 1 M HCl in ethanol (1 ml). The mixture was then filtered on a Whatman GF/A filter, washed extensively with ethanol, dried and counted in a scintillation counter. For gel filtration chromatography, 100 mg of an *E. coli* extract was applied onto a Sephacryl-300 (1.5 × 90 cm) column that was equilibrated with column buffer (50 mM Tris-chloride, pH 7.5, 50 mM NaCl, 1 mM MgCl₂, 1 mM DTT). Fractions (2.5 ml) were collected and 20 μl of each fraction was assayed for activity. The following column calibration standards were used: thyroglobulin (669K), alcohol dehydrogenase (150K), bovine serum albumin (66K), egg albumin (43K) and carbonic anhydrase (29K).

lipid donor was farnesyl pyrophosphate (FPP) (Fig. 1, lanes 3, 4, 7 and 8), instead of geranylgeranyl pyrophosphate (GGPP) (Fig. 1, lanes 1, 2, 5 and 6). Thus, like Ypt1p, Sec4p is geranylgeranylated in a Bet2p-dependent manner and it is not farnesylated. This result is consistent with earlier findings showing that the membrane attachment of Sec4p is defective in the *bet2* mutant¹².

The mammalian GGTase-II modifies the native form of Rab proteins that terminate in Cys-X-Cys or Cys-Cys sequences¹⁷⁻¹⁹. To determine whether the yeast enzyme behaves similarly, we used site-directed mutagenesis to modify the carboxy terminus of Ypt1p (Cys-Cys) to Cys-Ser-Cys (Ypt1p-CXC). Ypt1p-CXC is geranylgeranylated (Fig. 2c) and functions as the sole copy of Ypt1p in yeast (Y.J. and S.F.-N., unpublished observations), indicating that it is biologically active. As a control, we also constructed Ypt1p-CAAL. The CAAL (Cys-Thr-Ile-Leu) consensus sequence is recognized by Cdc43p, the β-subunit of the yeast GGTase-I (refs 5, 20). *In vitro*, the geranylgeranylation of

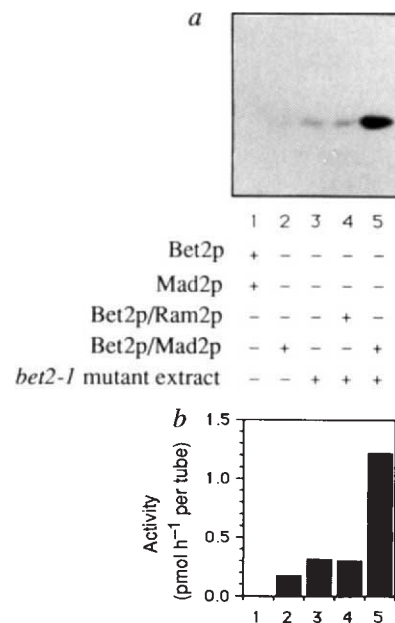


Ypt1p-CAAL was dependent on Cdc43p (not shown) and not Bet2p (Fig. 2d), whereas the geranylgeranylation of Ypt1p-CXC was greatly reduced when the prenylation assay was done with a *bet2* mutant extract (Fig. 2c). These findings indicate that, like the mammalian enzyme, the yeast GGTase-II geranylgeranylates proteins that terminate in the Cys-X-Cys sequence.

A *mad2* temperature-sensitive mutant (*mad2-2^{ts}*) displays the same phenotype as *bet2-1* (refs 12, 21 and 22). To test the hypothesis that Mad2p is a component of the yeast GGTase-II, we co-expressed Bet2p and Mad2p in *Escherichia coli* and then assayed soluble extracts for prenylation activity using a filter binding assay². Tritium-labelled lipid was trapped onto a Whatman filter when [³H]GGPP (Fig. 3a, lane 6), but not [³H]FFP (lane 7), served as the lipid donor. Furthermore, this activity was dependent on the co-expression of both proteins (Fig. 3a; compare lanes 1, 2 and 3 with lane 6) and no activity was detected if Bet2p was co-expressed with Ram2p (compare lanes 4 and 5 with lane 6), an α -subunit of the yeast farnesyltransferase and GGTase-I (ref. 9). When the soluble fraction (100,000g supernatant) of an *E. coli* extract, which co-expressed Bet2p and Mad2p, was chromatographed on a Sephacryl S-300 column, one peak of activity with $M_r \sim 70$ K was obtained (Fig. 3b). This is the predicted size of a Bet2p (36K)/Mad2p (35K) protein complex. Western blot analysis confirmed that a peak of Bet2p crossreacting material co-fractionated with this activity (not shown).

Because the filter assay described above cannot distinguish between the binding of lipid to the transferase and the covalent attachment of lipid to the acceptor, we also used a gel assay. Some geranylgeranyl was added onto Ypt1p when both proteins were co-expressed (Fig. 4, lane 2). In addition, no activity was detected when Bet2p and Mad2p were expressed separately and then mixed together (Fig. 4a, b, lane 1). The mammalian GGTase-II is composed of two of two components, A and B, that are both required for activity^{17,18}. Component B contains two polypeptides that are analogous to the α - and β -subunits of the CAAX prenyltransferase^{17,18}. Component A is the choroideraemia gene product²³ which also resembles GDI (GDP-dissociation inhibitor)^{23,24}, a protein that complexes with the geranylgeranylated form of small GTP-binding proteins²⁴. To determine whether yeast contains an activity that stimulates the transfer of geranylgeranyl onto Ypt1p, *bet2* mutant extracts, which have low transferase activity (Fig. 4a, b, lane 3), were mixed with bacterially expressed Bet2p/Mad2p and then assayed. We found that mutant extracts stimulated the transfer of geranylgeranyl onto Ypt1p in the presence of the Bet2p/Mad2p complex (Fig. 4a, b; compare lanes 2 and 3 with lane 5), but not when Bet2p was co-expressed with Ram2p (Fig. 4a, b; compare lanes 3 and 4 with lane 5). Furthermore, the addition of recombinant Bet2p to this extract did not stimulate transfer (not shown). Similar results were obtained when Sec4p

FIG. 4 Efficient transfer of geranylgeranyl onto Ypt1p requires a third component. a, Lane 1, *E. coli*-expressed Bet2p (50 μ g) mixed with *E. coli*-expressed Mad2p (50 μ g); lane 2, co-expressed Bet2p and Mad2p (50 μ g); lane 3, *bet2-1* mutant extract (50 μ g); lane 4, *bet2-1* mutant extract (50 μ g) mixed with co-expressed Bet2p and RAM2p (50 μ g); lane 5, *bet2-1* mutant extract (50 μ g) mixed with coexpressed Bet2p and Mad2p (50 μ g). b, Quantification of the results shown in a.



METHODS. The prenylation assay was done as before² with 0.4 μ M Ypt1p, 0.8 μ M [³H]GGPP and 50 μ g extract, and the reaction was stopped by addition of SDS sample buffer (final SDS concentration of 1%). For complementation assay, *E. coli* extracts were mixed with *bet2-1* mutant extracts and incubated at 30 °C for 10 min before the assay. About half the final sample was electrophoresed on a 12.5% SDS-polyacrylamide gel.

was used in place of Ypt1p in this assay (not shown). These data indicate that the efficient transfer of geranylgeranyl onto Ypt1p and Sec4p requires an additional activity which is provided by a *bet2* mutant extract.

We propose that Bet2p and Mad2p form a complex that is analogous to component B, whereas an unidentified subunit that behaves like component A binds to Ypt1p and Sec4p to stimulate the transfer of geranylgeranyl onto these small GTP-binding proteins. Yeast component A does not appear to be the same as GDI, because we have found that recombinant yeast GDI did not reconstitute prenyltransferase activity when it was incubated with co-expressed Bet2p/Mad2p (Y.J., M. Garrett, P. Novick and S.F.-N., unpublished observations). In support of the studies presented here, it was recently shown that the β -subunit of the mammalian GGTase-II is 52% identical to Bet2p (ref. 25). The α -subunit of this enzyme, which is significantly larger than Mad2p (60K versus 35K), is more homologous to Mad2p than Ram2p (ref. 25). Our data imply that this subunit is the mammalian homologue of Mad2p. The complementation assay that we have developed for component A in yeast will now enable us to purify and characterize this protein. □

Received 19 April; accepted 19 August 1993.

- Reiss, Y., Goldstein, J. L., Seabra, M. C., Casey, P. J. & Brown, M. S. *Cell* **62**, 81–88 (1990).
- Moores, S. L. et al. *J. Biol. Chem.* **266**, 14603–14610 (1991).
- Moomaw, J. F. & Casey, P. J. *J. Biol. Chem.* **267**, 17438–17443 (1992).
- Schaefer, W. R. et al. *Science* **249**, 1133–1139 (1990).
- Finogold, A. A. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4448–4452 (1991).
- Kohl, N. E. et al. *J. Biol. Chem.* **266**, 18884–18888 (1991).
- Reiss, Y., Stradley, S. J., Gierasch, L. M., Brown, M. S. & Goldstein, J. L. *Proc. natn. Acad. Sci. U.S.A.* **88**, 732–736 (1991).
- Seabra, M. C., Reiss, Y., Casey, P. J., Brown, M. S. & Goldstein, J. L. *Cell* **65**, 429–434 (1991).
- He, B. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11373–11377 (1991).
- Li, R. & Murray, A. *Cell* **66**, 519–531 (1991).
- Boguski, M. S., Murray, A. W. & Powers, S. *New Biol.* **4**, 1–4 (1992).
- Rossi, G., Jiang, Y., Newman, A. P. & Ferro-Novick, S. *Nature* **351**, 158–161 (1991).
- Segev, N., Mulholland, J. & Botstein, D. *Cell* **52**, 915–924 (1988).
- Bacon, R. A., Salminen, A., Ruohola, H., Novick, P. & Ferro-Novick, S. *J. Cell Biol.* **109**, 1015–1022 (1989).
- Baker, D., Wuesthube, L., Schekman, R., Botstein, D. & Segev, N. *Proc. natn. Acad. Sci. U.S.A.* **87**, 355–359 (1990).
- Salminen, A. & Novick, P. *Cell* **49**, 527–538 (1987).
- Seabra, M. C., Goldstein, J., Sudhof, T. C. & Brown, M. S. *J. Biol. Chem.* **267**, 14497–14503 (1992).

- Seabra, M. C., Brown, M. S., Slaughter, C. A., Sudhof, T. C. & Goldstein, J. L. *Cell* **70**, 1049–1057 (1992).
- Kinsella, B. & Maltse, W. A. *J. Biol. Chem.* **266**, 8540–8544 (1991).
- Mayer, M. L., Caplin, B. E. & Marshall, M. S. *J. Biol. Chem.* **267**, 20589–20593 (1992).
- Newman, A. & Ferro-Novick, S. *J. Cell Biol.* **105**, 1587–1594 (1987).
- Li, R., Havel, C., Watson, J. A. & Murray, A. W. *Nature* **366**, 82–84 (1993).
- Andres, D. A. et al. *Cell* **73**, 1091–1099 (1993).
- Araki, S., Kikuchi, A., Hata, Y., Isomura, M. & Takai, Y. *J. Biol. Chem.* **265**, 13007–13015 (1990).
- Armstrong, S. A., Seabra, M. C., Sudhof, T. C., Goldstein, J. L. & Brown, M. S. *J. Biol. Chem.* **268**, 12221–12229 (1993).
- Wagner, P. et al. *EMBO J.* **6**, 2373–2379 (1987).
- Walworth, N. C., Brennwald, P., Kabcenell, A. K., Garrett, M. & Novick, P. *Molec. cell. Biol.* **12**, 2017–2028 (1992).

ACKNOWLEDGEMENTS. We thank R. Li and A. Murray for the cloned *MAD2* gene, the discussion of unpublished results and comments on the manuscript, J. Goldstein for communicating unpublished results, M. Garrett for purified recombinant Sec4p and GDI, S. Powers and M. Garrett for plasmids, and H. Tan for photography. This work was supported by a Leslie H. Warner fellowship and an Arrighi L. and Anna G. Hull Cancer Research Award (G.R.), the National Cancer Institute (S.F.-N.) and the American Heart Association (S.F.-N.).

Yeast origin recognition complex is involved in DNA replication and transcriptional silencing

Gos Micklem^{*†}, Adele Rowley[‡], Janet Harwood[‡],
Kim Nasmyth[§] & John F. X. Diffley^{‡||}

^{*} MRC Laboratory of Molecular Biology, Hills Road,
Cambridge CB2 2QH, UK

[‡] ICRF Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD,
UK

[§] Research Institute of Molecular Pathology, Dr Bohr Gasse 7,
1030 Vienna, Austria

THE *HMR* E silencer represses transcription of silent mating-type genes in the budding yeast *Saccharomyces cerevisiae* and contains three redundant regulatory elements A, E and B (ref. 1). The A element contains the 11 base pair consensus sequence that is essential for the firing of DNA replication origins². A multisubunit protein called the origin recognition complex (ORC) binds specifically to this consensus sequence within yeast origins *in vitro*³ and *in vivo*⁴. We isolated mutants in A element-mediated silencing and report here that one of the genes we identified, *RRR1*, encodes ORC2, the 72K subunit of ORC. *RRR1/ORC2* is an essential gene, but the *rrr1-316* allele, which is viable, is defective in the replication of nuclear DNA and the maintenance of the 2- μ m episomal DNA. This is, to our knowledge, the first genetic evidence that ORC is involved in DNA replication and silencing.

The *HMR* E silencer blocks transcriptional activation from a *lacZ* reporter gene when placed between an upstream activating sequence (UAS) and its TATA box (Fig. 1). This does not seem to occur by the same mechanism as genuine silencing of *HMR* because the A element of *HMR* E is both necessary and sufficient for efficient transcriptional blocking in this context (Fig. 1b, c). We exploited this fact to identify genes whose products might interact with the autonomously replicating sequence (ARS) consensus sequence. We isolated 110 mutants that expressed *LacZ* from a blocked construct; they defined five main complementation groups (see legend to Fig. 1).

Deletion of both the A and E or A and B elements cause *HMRa* to be derepressed, whereas mutation of any single element has little (E) or no (A and B) effect. Mutations from only two complementation groups, called *RRR1* and *RRR2* (redundant regulator of *HMR*) caused extensive derepressed transcription of an *HMRa* locus lacking its silencer E element (Fig. 2a). Only *rrr1* mutants have been further analysed. Transcription from an *HMRa* locus in which the silencer lacked the B element was not affected by the *rrr1-316* mutation to the same extent as when the silencer lacked the E element, suggesting that the A element retains some function, consistent with the fact that this allele is viable whereas a null allele is lethal (see below).

In addition to defects in transcriptional repression, three out of four *rrr1* alleles show a severe growth defect (Fig. 2b and data not shown). Flow cytometry (Fig. 2c) showed that asynchronous cultures of *rrr1-316* mutants contain fewer cells with a 1N DNA content and more cells with DNA contents between 1N and 2N, a pattern similar to that in mutants deleted for the B-type cyclin, *CLB5*, implicated in the initiation of DNA replication^{5,6}. Together with the large increase in generation time, these results suggest that S phase is extended in *rrr1* mutant cells. The endogenous 2- μ m plasmid, in which replication occurs by a mechanism very similar to that of chromosomal origins⁷⁻¹⁰, is normally

present at 60–100 copies per haploid cell¹¹. This copy number is reduced more than 10-fold in *rrr1-316* mutants (Fig. 2d), consistent with the notion that *rrr1-316* mutants are defective in DNA replication. This does not rule out a role for *RRR1* in other aspects of 2- μ m function such as amplification during S phase.

The *RRR1* gene, which was cloned by complementation of the growth defect of the *rrr1-316* mutant (see legend to Fig. 3), also complements the defect in transcriptional silencing in an *HMR* E element mutant (Fig. 3b). This gene encodes a polypeptide of 71.2K with no significant similarity to any known protein. One copy of this open reading frame was disrupted in a diploid strain with a *URA3* marker (Fig. 3a) and the products of meiosis were examined. In no case were more than two viable spores produced per tetrad and in no case were the viable spore products *Ura*⁺. These results indicate that the *RRR1* gene is essential for proliferation.

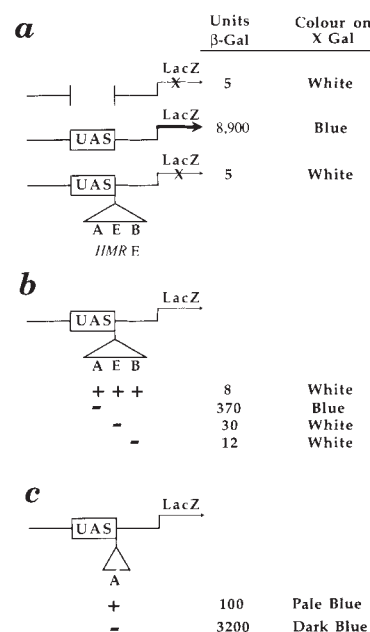


FIG. 1 Blocking of transcriptional activation by *HMR* E. **a**, The *HMR* E silencer blocks transcriptional activation by the *CYC1* UAS. Cells carrying plasmids containing the indicated *lacZ* reporter constructs were assayed for β -galactosidase activity by both a qualitative colony colour assay using X-gal and a semi-quantitative assay of cell extracts using ortho-nitrophenyl- β -D-galactopyranoside (ONPG) as the colorogenic substrates. **b**, Mutations in either the E or B elements have only a small effect on transcriptional blocking whereas mutation of the A element has a severe effect on blocking. **c**, The A element alone is sufficient for efficient blocking and mutation of this sequence eliminates blocking. **METHODS.** The *DraI*-*AluI* fragments containing 138 base pairs (bp) covering *HMR* E wild type and A, E or B lesions were cloned into blunt-ended, *XhoI*-cut pLGA312 (ref. 16). Lesions in the A, E and B elements were those used in ref. 1. The A element constructs carry a 55 bp *DraI*-*XhoI* fragment from the E element linker substitution silencer 331–324 (ref. 1) cloned into the *XhoI* site of pLGA312 (ref. 16). β -Galactosidase activity was assayed qualitatively on filters using X-gal and quantitatively using ONPG as previously described^{17,18}. Units are defined as absorbance at 420 nm (A_{420}) (units per mg protein per min). Four strains were mutagenized with ethyl methanesulphonate (EMS) to 10–50% survival, plated on YEPD plates at 25 °C, then transferred to nitrocellulose, incubated at 37 °C for 4 h and assayed for β -galactosidase activity on X-gal. The four strains used were derived from W303-1b (ref. 19; *MATa his3-11, 15 trp1-1 ade2-1 leu2-3,112 ura3 can1-100*) and contained either wild type or E or B element mutations¹. The 55 bp A element blocking construct described above was integrated at *URA3* in the mutant silencer backgrounds described in ref. 1 after removing a *HindIII* fragment removing 2- μ m sequences. Other details of analysis of mutants and cloning of the *RRR1* gene can be found in ref. 20.

|| To whom correspondence should be addressed.

† Present address: Sanger Centre, Hinxton, Cambridge CB10 1RQ, UK.

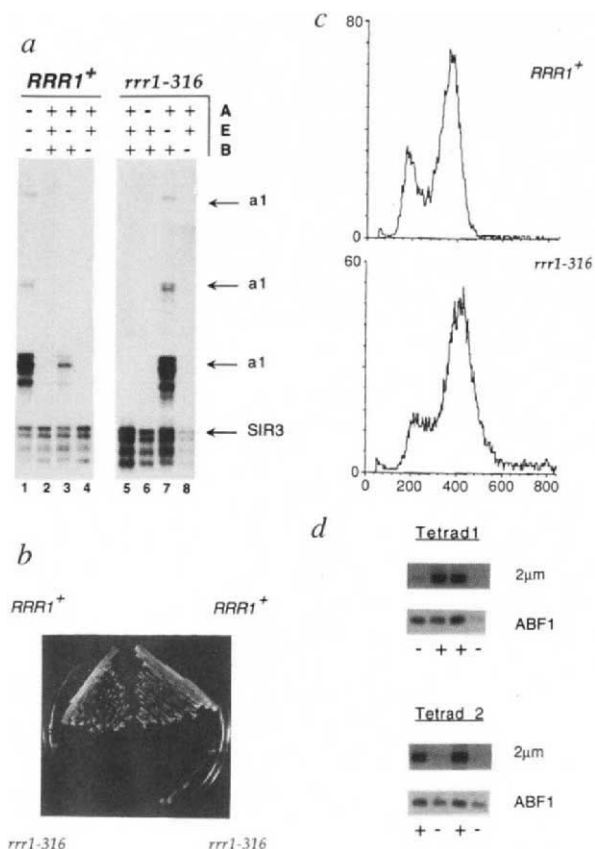


FIG. 2 *rrr1* mutants are defective in transcriptional silencing and DNA replication. **a**, *rrr1* mutants are defective in transcriptional silencing by a non-redundant silencer. An *RRR1* strain and an *rrr1-316* strain were crossed into strains in which *HMR* silencing is under the control of the indicated wild-type and mutant silencers and the level of *HMRa* a1 transcript was determined by S1 nuclease protection²¹ using the transcript from the *SIR3* gene as an internal control²². **b**, *rrr1-316* mutants show an extreme growth defect. A strain containing the *rrr1-316* allele was crossed to a wild-type strain (YGM185), sporulated and tetrads were dissected. This figure shows 2 days growth after streaking the products of a representative tetrad on YEPD medium at 30 °C. **c**, DNA flow cytometry of asynchronous cultures. log-phase YEPD cultures were sonicated, fixed and stained with propidium iodide. DNA was quantified as described⁵ on a Becton-Dickson FACScan. **d**, The 2-μm copy number is abnormally low in *rrr1* mutants. Two tetrads derived from a cross between a *rrr1-316* mutant and a wild-type strain (YGM185) were analysed for 2-μm copy number by DNA blot hybridization using a probe from 2-μm and from the chromosomal *ABF1* gene, which serves as an internal control. Products of the cross were *RRR1*(+) and *rrr1-316*(-). In all cases, *rrr1* mutant phenotypes co-segregated in crosses with wild-type strains.

METHODS. S1 nuclease protection analysis of a1 was done on total RNA with *SIR3* transcripts used as an internal control as described^{1,23}. The two larger a1 protected species are immature intron-containing mRNAs²⁴. In **d**, DNA was digested with *NcoI*, which linearizes the 2-μm circle (6.3 kilobases (kb)) and generates a 1.6 kb fragment within the *ABF1* gene²⁵. The blot in **d** was analysed on a Molecular Dynamics Phosphorimager to quantify the levels of 2-μm DNA and *ABF1*, indicating a 10.9-fold reduction in the copy number of 2-μm DNA after normalizing to the internal *ABF1* control.

ORC is composed of subunits of 120, 72, 62, 56, 53 and 50K (ref. 3 and Fig. 4b) and five independent polyclonal antisera raised against a truncated *RRR1* protein expressed in *Escherichia coli* specifically crossreacted with the 72K subunit of purified ORC (Fig. 4c and co-fractionation data not shown). To provide further evidence that *RRR1* encodes the 72K subunit of ORC, the *RRR1* gene of a wild-type cell was tagged at its C

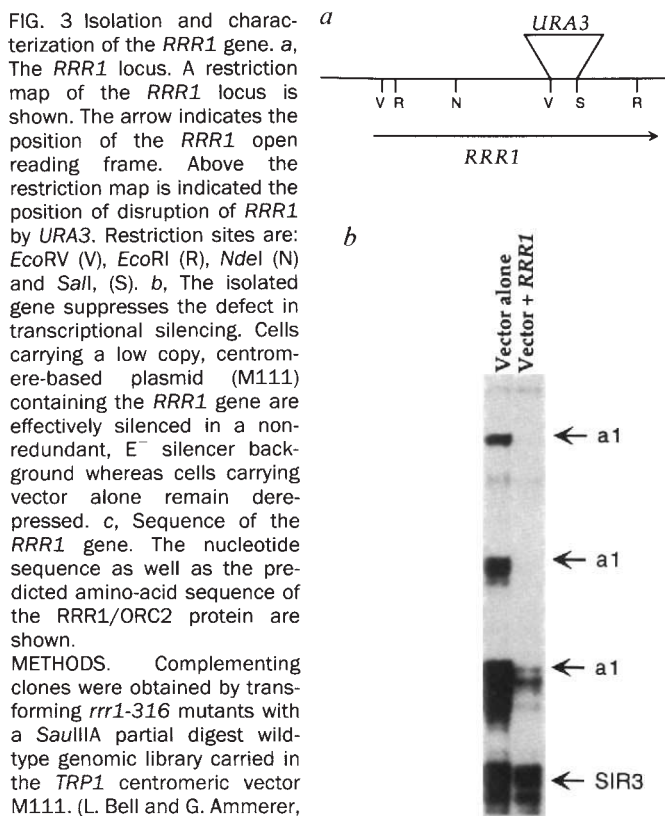


FIG. 3 Isolation and characterization of the *RRR1* gene. **a**, The *RRR1* locus. A restriction map of the *RRR1* locus is shown. The arrow indicates the position of the *RRR1* open reading frame. Above the restriction map is indicated the position of disruption of *RRR1* by *URA3*. Restriction sites are: *EcoRV* (V), *EcoRI* (R), *NdeI* (N) and *Sall*, (S). **b**, The isolated gene suppresses the defect in transcriptional silencing. Cells carrying a low copy, centromere-based plasmid (M111) containing the *RRR1* gene are effectively silenced in a non-redundant, E⁻ silencer background whereas cells carrying vector alone remain derepressed. **c**, Sequence of the *RRR1* gene. The nucleotide sequence as well as the predicted amino-acid sequence of the *RRR1*/ORC2 protein are shown.

METHODS. Complementary clones were obtained by transforming *rrr1-316* mutants with a *SauIII*A partial digest wild-type genomic library carried in the *TRP1* centromeric vector M111. (L. Bell and G. Ammerer, personal communication), and screening for clones in which the growth defect was complemented. Three overlapping clones were thus obtained. The region containing the *RRR1* gene was sequenced from both single-stranded M13 and double-stranded plasmids. Both strands were sequenced at least twice. The minimal region of overlap began at a *SauIII*A site corresponding to amino acids 91–92 and the sequence was submitted to EMBL with accession number Z21817. The complete sequence has now been submitted to EMBL. All integrations were confirmed by agarose gel electrophoresis and DNA blot hybridization using probes from the *RRR1* gene. A marker gene (*TRP1*) integrated into a *RRR1* strain adjacent to the cloned gene was closely linked to *RRR1* in crosses with an *rrr1* mutant (data not shown) demonstrating that the cloned gene is not an unlinked suppressor. Thirteen tetrads from a heterozygote in which one copy of the *RRR1* gene was disrupted as indicated in **a** were dissected on YEPD plates. Twelve tetrads produced two viable spores and one tetrad produced one viable spore. None of the viable spores were *URA*⁺.

terminus¹² with the 10 amino-acid c-myc epitope recognized by the 9E10 monoclonal antibody¹³ (Fig. 4a). When ORC was purified to near homogeneity from strains carrying either the tagged (+) or untagged (–) genes, only ORC2 from the tagged strain was recognized by the 9E10 antibody (Fig. 4d); note also that the tagged protein migrated slightly slower than the wild-type protein (Fig. 4b, c), consistent with the addition of the 1.5K myc epitope.

Our data indicate that the *RRR1* gene encodes ORC2 and provide the first genetic evidence that ORC is involved in both DNA replication and transcriptional silencing. It will be interesting to determine whether factors such as ORC2 are conserved among eukaryotic organisms. The mechanism by which ORC is controlled during the cell cycle remains to be elucidated. *In vivo* footprinting studies indicate that ORC is not present transiently, but rather remains bound at replication origins throughout most or all of the cell cycle (ref. 4 and J.D., S. Dowell and A.R., manuscript in preparation). This might be consistent with the role of ORC in transcriptional silencing described here, as the *HM* loci must remain transcriptionally silent throughout the cell cycle. The initiation of DNA replication, however, is limited to S phase and initiation complexes must somehow be regulated accordingly. In this regard, it is interesting that ORC2 contains

c

TAGCAATAGCGTCACTAGCTGCCGTGACATTAACCTTGCTGTGGCACTTTATATGTAATGAACCATCTTTCAATGGATCAT
 AAGAAATAGTGTGTTAAAGGCCAAATATCCATGCATAAATATGACCTTATTCGGTAATGTGATATGGATCAGCTAGTACCC
 AATTTCTAGTCTAGCAAAATCGGAAATTTTTCAGAACACCCACTCAGCGCATCATTTAGGTTGGAATGACAAATAGTAAAGCA
 GAATTTGTTATTTCTCAAAATGTGTAAGATTTAATAAGAAATAGGAACCACTTTAAATTAAGACAAAGATGAATATATTAGCT
 GAAATGTATTGTAATATGATCATTTGCTATATCTTTAAACAAGTTTGTAGTACTGCGCAATTCGCAATAAC

1 MET Leu Asn Gly Glu Asp Phe Val Glu His Asn Asp Ile Leu Ser Ser Pro Ala Lys Ser
 21 Arg Asn Val Thr Pro Lys Arg Val Asp Pro His Gly Glu Arg Gln Leu Arg Arg Ile His
 41 Ser Ser Lys Lys Asn Leu Leu Glu Arg Ile Ser Leu Val Gly Asn Glu Arg Lys Asn Thr
 61 Ser Pro Ser Lys Ala Leu Lys Pro Lys Thr Pro Ser Lys Ala Pro Ser Lys Lys Ser Lys
 81 Pro Arg Lys Ile Gln Glu Glu Leu Thr Asp Arg Ile Lys Lys Asp Glu Lys Asp Thr Ile
 101 Ser Ser Lys Lys Lys Arg Lys Leu Asp Lys Asp Thr Ser Gly Ser Ala Ser Glu Phe Lys
 121 Lys Thr Ser Asn Asn Lys Gln Val Met Glu Lys Thr Gly Ile Lys Glu Lys Arg Glu Arg
 141 Glu Lys Ile Gln Val Ala Thr Thr Thr Tyr Glu Asp Asn Val Thr Pro Gln Thr Asp Asp
 161 Asn Phe Val Ser Met Ala Thr Pro Ser Lys Lys Gln Ile Ile Met Asn Asn Leu Lys Glu
 181 Thr Thr Asn His Asp Phe Thr Ser Pro Leu Lys Lys Gln Ile Ile Met Asn Asn Leu Lys Glu
 201 Tyr Lys Asp Ser Thr Ser Pro Gly Lys Leu Thr Leu Ser Arg Asn Phe Thr Pro Thr Pro
 221 Val Pro Lys Asn Lys Lys Leu Tyr Gln Thr Lys Ser Glu Thr Lys Ser Ala Ser Ser Phe Leu
 241 Asp Thr Phe Glu Gly Tyr Phe Asp Gln Arg Lys Ile Val Arg Thr Asn Ala Lys Ser Arg
 261 His Thr Met Ser Met Ala Pro Asp Val Glu Glu Phe Ser Leu Val Ser Asn Phe
 281 Phe Asn Glu Asn Phe Gln Lys Arg Pro Arg Gln Lys Leu Phe Glu Ile Gln Lys Lys Met
 301 Phe Gln Thr Trr Phe Gln Lys Arg Pro Arg Gln Lys Leu Phe Glu Ile Gln Lys Lys Met
 321 Ser Lys Arg Asn Phe Leu Glu Glu Phe Ala Ile Asp Tyr Leu Ser Pro Lys Ile Ala Tyr
 341 Ser Gln Leu Ala Tyr Glu Asn Glu Leu Gln Gln Asn Lys Pro Val Asn Ser Ile Pro Cys
 361 Leu Leu Asn Gly Tyr Asn Pro Ser Cys Asn Tyr Arg Asp Val Phe Lys Glu Ile Thr
 381 Asp Leu Leu Val Pro Ala Glu Leu Thr Arg Ser Glu Thr Lys Tyr Trp Gly Asn His Val
 401 Ile Leu Gln Ile Gln Lys Met Ile Asp Phe Tyr Lys Asn Gln Pro Leu Asp Ile Lys Leu
 421 Ile CTT GTA GTG CAT AAT CTG GAT GGT CTT AGC ATA AGG AAA AAC CAT TTT CAG ACG ATG
 441 Leu Ser Phe Leu Ser Val Ile Arg Gln Ile Ala Ile Val Ala Ser Thr Asp His Ile Tyr
 461 Ala Pro Leu Leu Trp Asp Asn Met Lys Ala Asn Tyr Asn Phe Val Phe His Asn Ile
 481 Ser Asn Phe Glu Pro Ser Thr Val Glu Ser Thr Phe Gln Asp Val Met Lys Met Gly Lys
 501 Ser Asp Thr Ser Ser Gly Ala Glu Gly Ala Lys Tyr Val Leu Gln Ser Leu Thr Val Asn
 521 Ser Lys Met Tyr Lys Leu Thr Gln Met Gln Lys Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
 541 Ala Asn Thr Gly Pro Lys Arg Gly Thr Gln Arg Thr Gly Val Glu Leu Lys Leu Phe Asn
 561 His Lys Lys Ala Ala Lys Phe Ile Ala Ser Asn Glu Ile Ala Leu Arg Ser Met Leu Arg
 581 Glu Phe Ile Glu His Lys Met Ala Asn Ile Thr Lys Asn Asn Ser Gly Met Glu Ile Ile
 601 Trp Val Pro Tyr Thr Tyr Ala Glu Leu Glu Lys Leu Leu Lys Thr Val Leu Asn Thr Leu
 TAAATGTATACATATACCAACAATTTGTAATAGTACTAGCTTGTGCTTTGCCATAACCAAC

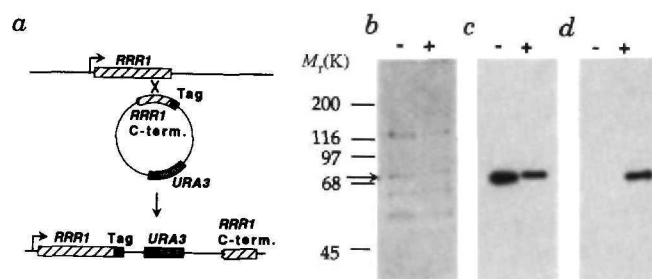


FIG. 4 RRR1 encodes the 72K subunit of ORC. **a**, Construction of a tagged *RRR1* gene. The 3' end of the *RRR1* open reading frame (*RRR1* C term.) was fused in-frame to sequences encoding a 1.5K epitope (Tag, amino-acid sequence EQLISEDLRS; single-letter code) specifically recognized by the 9E10 monoclonal antibody¹³. This tagged *RRR1* C-terminal gene fragment was integrated at the *RRR1* locus generating one complete tagged copy under the control of its natural promoter and one N-terminally truncated (and non-functional) copy of the *RRR1* gene. **b**, Subunit structure of ORC purified from untagged (–) and tagged (+) yeast visualized after SDS–PAGE by Coomassie blue staining. **c**, The 72K subunits of ORC from both untagged and tagged strains are recognized by polyclonal antiserum directed against the *RRR1* gene product. **d**, Tagged ORC2 is specifically recognized by the 9E10 monoclonal antibody.

METHODS. ORC was purified from strain PY26 (*MATa ura3-52 leu 2-3,112 trp1Δ prb1-1122 prc1-407 pep4-3 NUC1::LEU2*) and PY26 containing tagged *RRR1* using DNaseI footprinting assay and a protocol (A.R. and J.F.X.D., manuscript in preparation) based on that previously described³. About 500 ng per lane of ORC from each preparation was subjected to electrophoresis through 10% polyacrylamide gels containing SDS in triplicate. One set was stained with Coomassie blue (Fig. 4b) whereas the other two were used in western blot analysis after transfer to nitrocellulose (c, d). Identical blots were probed with either rat polyclonal antiserum (c) or the c-myc-specific antibody 9E10 (d). Rat polyclonal antiserum was used at a 1:500 dilution and affinity purified 9E10 was used at 5 µg ml^{–1}.

at least five potential *CDC28* kinase phosphorylation sites^{14,15} (Fig. 3c) because activation of this kinase by the B-type cyclins *CLB5* and *CLB6* has been implicated in the initiation of S phase. □

Received 3 September; accepted 11 October 1993.

- Brand, A. H., Micklem, G. & Nasmyth, K. *Cell* **51**, 701–719 (1987).
- Campbell, J. L. & Newlon, C. S. in *The Molecular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis and Energetics* (eds Broach, J. R., Pringle, J. R. & Jones, E. W.) 41–146 (Cold Spring Harbor Laboratory Press, New York, 1991).
- Bell, S.P. & Stillman, B. *Nature* **357**, 128–134 (1992).
- Diffey, J. F. X. & Cocker, J. H. *Nature* **357**, 169–172 (1992).
- Epstein, C. B. & Cross, F. R. *Genes Dev.* **6**, 1695–1706 (1992).
- Schwob, E. & Nasmyth, K. *Genes Dev.* **7**, 1160–1175 (1993).
- Livingston, D. M. & Kupfer, D. M. *J. molec. Biol.* **116**, 249–260 (1977).
- Zakian, V. A., Brewer, B. J. & Fangman, W. L. *Cell* **17**, 923–934 (1979).
- Brewer, B. J. & Fangman, W. L. *Cell* **51**, 463–71 (1987).
- Huberman, J. A., Spotila, L. D., Nawotka, K. A., El Assouli, S. M. & Davis, L. R. *Cell* **51**, 473–81 (1987).
- Broach, J. R. & Volker, F. C. in *The Molecular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis and Energetics* (eds Broach, J. R., Pringle, J. R. & Jones, E. W.) 297–332 (Cold Spring Harbor Laboratory Press, New York, 1991).
- Kolodziej, P. A. & Young, R. A. in *Meth. Enzym.* **194**, 508–519 (1991).
- Evan, G., Lewis, G., Ramsay, G. & Bishop, J. M. *Molec. cell. Biol.* **5**, 3610–3616 (1985).
- Langan, T. A. et al. *Molec. cell. Biol.* **9**, 3860–3868 (1989).
- Shenoy, S. et al. *Cell* **57**, 763–774 (1989).
- Guarente, L. & Mason, T. *Cell* **32**, 1279–1286 (1983).
- Breeden, L. & Nasmyth, K. *Cold Spring Harb. Symp. quant. Biol.* **50**, 643–650 (1985).
- Breeden, L. & Nasmyth, K. *Cell* **48**, 389–397 (1987).
- Thomas, B. T. & Rothstein, R. *Cell* **56**, 619–630 (1989).
- Micklem, T. G. thesis, Cambridge Univ. (1989).
- Nasmyth, K. *Nature* **302**, 670–676 (1983).
- Shore, D., Squire, M. & Nasmyth, K. A. *EMBO J.* **3**, 2817–2823 (1984).
- Brand, A. H., Breeden, L., Abraham, J., Sternglanz, R. & Nasmyth, K. *Cell* **41**, 41–48 (1985).
- Miller, A. M. *EMBO J.* **3**, 1061–1065 (1984).
- Diffey, J. F. X. & Stillman, B. *Science* **246**, 1034–1038 (1989).

ACKNOWLEDGEMENTS. G.M. and A.R. should be considered as equal first authors. We thank our colleagues at ICRF, Clare Hall and the LMB, Cambridge for discussions, E. Schwob for flow cytometry, D. Micklem for *E. coli* expression of *RRR1*, P. Wright at LMB for animal care, I. Goldsmith for oligonucleotide synthesis, G. Evan for peptide synthesis and monoclonal antibody and A. Brand for strains and plasmids.

Huntington's is still holding out

***In situ* expression studies show that the selective pathology of Huntington's disease is not explained by a simple pattern of gene expression in the brain.**

HUNTINGTON'S disease is a disorder to be reckoned with. The causative gene managed to remain hidden for a full ten years of intensive international gene searching after linkage was established, and now the fear that finding it might have been the easy bit seems to be being borne out. As Francis Collins and his colleagues report in this month's *Nature Genetics*¹, the pattern of gene expression does nothing to clarify the intriguing specificity of the disease pathology.

In 1872 George Huntington² described an inherited chorea that he, his father and his grandfather had all remarked upon. But the description was not widely read, and it had to wait 20 years before W. Osler once more drew attention to the heritable chorea³ which is also characterized by progressive intellectual deterioration and severe depression. At this time Osler remarked that he had attempted to investigate the original family described by Huntington, only to be told that the patients could not be seen because of their sensitivity about their condition.

For many years, remarkably little progress was made in describing the condition, although in 1932 the original families with Huntington's disease (HD) were traced back to a pair of English brothers⁴ and hence the stability of the disease was established. In the 1950s and 1960s, several groups reported on large studies documenting the huge variation in age of onset (typically between the third to fifth decade, although some patients develop signs in the first decade and others are symptom-free for more than 60 years) and death (usually ten years after onset of symptoms). In the 1970s, some progress was made in explaining these variations when it was noted that the juvenile form, with a very early onset, was significantly more frequent in those who had inherited the defective gene from their father. As this theme continued to develop, the possibility of chromosomal imprinting became clear.

In the 1980s, molecular genetic linkage

studies gained ground as the method of choice for identifying wayward genes. After a number of false starts, an anonymous marker (G8, mapped to chromosome 4; ref. 5) was found to be linked to the disease in a massive Venezuelan kindred. (This huge pedigree also provided access to homozygote HD patients who, on the basis that they had a phenotype indistinguishable from heterozygote patients, proved that HD showed complete dominance.)

The first, albeit rudimentary, localization of the HD gene was expected to herald great things. The devastating nature of the disease had become widely understood and mapping the HD gene to chromosome 4 was the first really tangible step towards isolating, understanding and presumably treating or even curing the condition. But despite the close attention of the medical and research communities, progress was slow, with little or no significant advance either in understanding the basic defect, the selective neuronal loss seen in the striatum, or in the treatment of patients. HD continued to prove frustrating and puzzling. New and ever-closer markers were described, and over a period of ten years the gene locus was gradually refined until just the very tip of the long arm of chromosome 4 was in the spotlight. Elaborate and powerful molecular techniques were used to scour the region and candidate genes were picked up and discarded.

In March of this year the breakthrough came when the Huntington's Disease Collaborative Research Group described a large (210-kilobase) gene that included a polymorphic trinucleotide (CAG)*n* repeat which is expanded in HD chromosomes⁶. Early analysis (since confirmed in larger studies) showed that in normal chromosomes the repeat was always 34 units or less, whereas in HD chromosomes it was 42 or more. The excitement of finding the HD gene was all the greater with the discovery of another unstable triplet repeat giving rise to yet another neurological disorder⁷.

Unfortunately the ten kilobases of HD transcript (IT15) reveal a sequence unrelated to any previously described, and hence predict a protein of unknown function, making comparisons impossible. Strong *et al.* therefore tried a different tack and postulated that the selective neuronal loss in brains of HD patients might reflect selective expression of

IT15. Rat and human brain sections were both tested using RNA *in situ* hybridization with riboprobes directed against the human IT15 transcript or the rat homologue. Looking in sections taken from normal and diseased brains, they found that in both rat (normal) and human (normal and HD) sections the HD gene appears to be expressed in neurons throughout the brain. Although some variation in expression levels was seen, the striatum (the area of severe neuronal loss in HD brains) did not show unusually high expression, and in the human sections there was no difference between the HD and control brain. HD messenger RNA was also found in non-neural tissues (pancreas, colon, liver and sperm), although at slightly reduced levels compared with brain. From a cursory glance at the generalized brain and non-neural expression (and bearing in mind that no HD pathology has been described in extraneural tissues), it is not clear that HD pathology is a function of IT15 expression. These results have been confirmed by a second group⁸.

Strong *et al.* point out that as HD appears to act in a truly dominant manner a gain-of-function mutation is a plausible disease mechanism and their results do not support this idea, at least not in any straightforward fashion. They also point out that they only looked at adult brain tissue and a transient overexpression of IT15 could occur, or there may be something about neuronal cells of the striatum that makes them particularly sensitive to IT15. No doubt many other explanations will be suggested, which may demand many years of investigation.

Shortly after the gene linkage was published, Peter Little wrote a *Nature News & Views* article⁹ on why the search for the HD gene had been so protracted. He concluded "Isolating the gene is, in truth, only the end of the beginning."

Adrian J. Iverson

Adrian J. Iverson is the Assistant Editor of *Nature Genetics*.

Also in this month's *Nature Genetics*: a beautiful explanation of the lost haemophilia A mutations; protein zero and peripheral myelin protein 22 mutations in Dejerine-Sottas disease; chromosomal background affects cystic fibrosis phenotypes; and mapping Stargardt's disease.

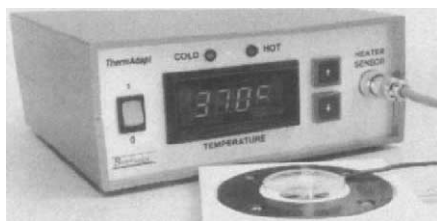
1. Strong, T. V. *et al.* *Nature Genet.* **5**, 259-265 (1993).
2. Huntington, G. *Med. Surg. Reporter* **26**, 317-321 (1872).
3. Osler, W. *J. Nerv. Ment. Dis.* **18**, 97-111 (1893).
4. Vessie, P. R. *J. Nerv. Ment. Dis.* **76**, 553-573 (1932).
5. Gusella, J. F. *et al.* *Am. J. Hum. Genet.* **36**, 139 abstr. (1984).
6. The Huntington's Disease Collaborative Research Group *Cell* **72**, 971-983 (1993).
7. Brook, J. D. *Nature Genet.* **3**, 279-281 (1993).
8. Li, S.H. *et al.* *Neuron* (in the press).
9. Little, P. *Nature* **362**, 408-409 (1993).

New tools for neuroscientists

News from the marketplace this week features a fluorescence imaging system for recording intracellular ion concentrations, a multichannel tissue bath stimulation system and flow-through sensors for organ perfusion studies.

INACTIN (thiobutabarbital sodium salt), a **long-lasting rodent anaesthetic**, is now available from RBI through its UK distributor, Semat Technical (*Reader Service No. 100*). Because of its rapid onset and duration of action, Inactin should be of interest to physiologists using small animals. It is most useful for long experiments, or for experiments requiring minimal anaesthetic disruption of smooth muscle tone and/or renal output. It is long lasting; Semat says that more than 8 h of anaesthesia can be achieved with a single intraperitoneal dose; pentobarbital, by comparison, is said to give only 1 h of anaesthesia. Inactin is supplied in serum vials for reconstitution.

The new **microincubation chamber** from Biophysica Technologies provides viewing, access and thermoregulation for



Thermoregulation from -4°C to 50°C with the new microincubation chamber from Biophysica Technologies.

tissues or cultured cells on inverted microscopes (*Reader Service No. 101*). The system supports thermoregulation from -4°C to 50°C (with a stated precision of $\pm 0.2^{\circ}\text{C}$), micromanipulation, perfusion, atmospheric control and long-term microscopic studies under sterile conditions. It can also be used with fluorescent optical indicators such as fura-2, indo-1 and BCECF for measuring intracellular calcium and pH. Its optional ultrathin (0.07 mm) coverslip well bottom enhances use of short-working-distance objectives on inverted microscopes. The coverslips also fit in standard 35 mm culture dishes.

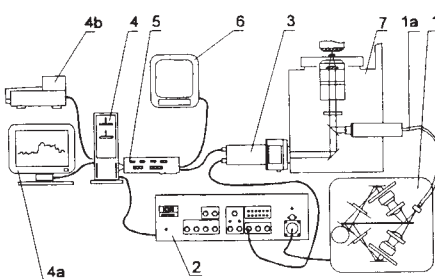
New from Oncor Instrument Systems, formerly American Innovision, is the Videometric 150 PC-based, **real-time image analysis system** (*Reader Service No. 102*). Among the system's features is an HSI colour analysis technology for the accurate identification of cell or tissue types for various applications. Some of the stated applications include the analysis and quantification of immunohistochemistry (positivity level or negative versus positive), H and E stained tissue morphology, fluorescent marker iden-

tification and analysis, *in situ* hybridization autoradiography studies and general morphological measurements. Additionally, a large macro-program library exists for a wide variety of applications. The Videometric 150 provides the user with several benefits, including fast true-colour, real-time analysis, an expanded macro language providing greater flexibility, an x,y,z microscope stage for automation, low-light RGB colour cameras for fluorescent applications.

■ Ion recording and analysis

BDS Ratio++ version 2.0, released in August by Biological Detection Systems, features **high-speed image ratioing**, time-lapse imaging and video-enhanced contrast (*Reader Service No. 103*). This new application package for the Apple Macintosh utilizes an intuitive user interface called VOI (Visual Object Interface). BDS RATIO++ performs high-speed image ratioing for calcium, pH and other ions, as well as other 'live' cell analyses. Additional features include multiple regions of interest, real-time trend charting, video-enhanced contrast, temporal feature tracking and high-speed cooled CCD frame transfer support.

IonOptix has developed a modular **dual-excitation fluorescence imaging system** for recording and analysis of intracellular ion concentrations of Ca^{2+} , Mg^{2+} , K^{+} , Na^{+} , Cl^{-} and pH in isolated cell experiments (*Reader Service No. 104*). Fluorescence of indicator-loaded cells is observed by presenting excitation light of a selected wavelength and examining the emission image. Excitation light is provided by IonOptix's patented chopper-based xenon light source (1), equipped with a computer-controlled electronic shutter and stepper motor-driven filter wheel in each excitation path to permit recording from various fluorescent probes. The company's fluorescence interface (2) connects the system software to the hard-

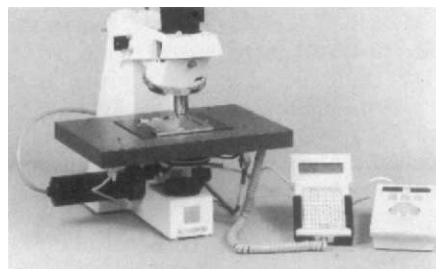


Fluorescent imaging ion quantitation with IonOptix's dual-excitation fluorescence imaging system.

ware elements and also provides the external triggering functions. An optical light guide (1a) delivers the excitation light to the fluorescence microscope (7). Emission images are recorded by an intensified CCD camera (3), coupled to the microscope, and are digitized by a high-performance frame grabber board (5). Both online and ratio/calcium images are viewed on a multi-input analog RGB monitor (6). Software for DOS-Windows environments is designed to provide easy access and operation of the hardware and off-line trace and image analysis. Major features of the system include: dual-excitation fluorescence imaging with video-rate wavelength alternation, sampling rates of 30 ratio images per second, automatic dual-dye recording, simultaneous analog-to-digital conversion of membrane potential, current, cell length and tension, and the capability to be upgraded for photon counting.

■ Microscope add-ons

Designed to replace manual stages in standard upright and inverted microscopes, the new SmartStage **motorized microscope stage** from Cell Robotics is said to make



Cell Robotics provides a motorized means with the SmartStage microscope stage.

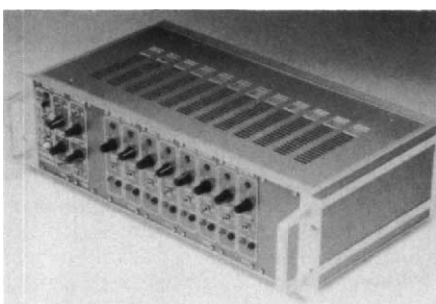
accurate scanning of slides and samples quick, simple and easily repeatable (*Reader Service No. 105*). The SmartStage features a 75 mm $\times$ 50-mm range of motion along the x-axis and y-axis, respectively, and can accommodate three standard microscope slides at the same time. Petri dishes, culture flasks and other devices may also be used. An optional motorized focus adjust allows full three-axis control of all parameters using a single trackball interface. A microprocessor-based controller, housed within the x-y stage, connects to the trackball device and a handheld RS232 terminal (or a peripheral computer with a standard communications software package may be substituted). The SmartStage responds to simple text commands, so no special host software is required. In addition, SmartStage allows use of either absolute or relative

coordinates for x - y motion. Any reference location can be selected by moving to that point using positioning commands from the terminal or by using the trackball and then entering a single character command. Scanning matrices are defined by the user, oriented by column or row. The size of the fields within the matrix are also user-defined, after which the scanning process may be manually or automatically controlled.

Carl Zeiss has developed ICS (Infinity Colour-Corrected) objectives for Ca^{2+} measurements in living cells (*Reader Service No. 106*). These objectives are intended for use in microscopy for the physiological examination of cells, particularly when the Ca^{2+} probe fura-2 is used, displaying a transmission of over 70 per cent at a wavelength of 340 nm. The series of 'Fluar' objectives are also suitable for transmitted-light techniques, such as brightfield, phase contrast and differential interference contrast. Together with the short, direct fluorescence illuminating and imaging paths in the inverted and upright Zeiss research microscopes Axioskop, Axioplan and Axiovert and the LSM410 inverted confocal laser scanning microscope, these ICS objectives create conditions for optimum transmission efficiency: in all light-microscopic measurements of Ca^{2+} transients in living cell and tissue material using photometry, image analysis and confocal microscopy. The range of Fluar objectives includes the Fluar $\times 10/0.5$, $\times 40/1.3$ oil and $\times 100/1.3$ oil.

■ Items of hardware

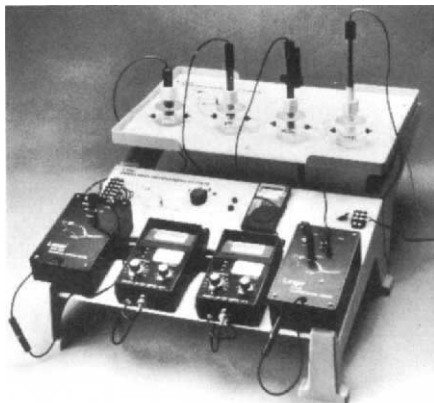
The Digitimer Multistim multichannel tissue bath stimulation system from Medical Systems has been designed to facilitate multiple isolated tissue bath stimulation, with flexible control of stimulation voltage and timing for each bath (*Reader Service No. 107*). The system can be used for isolated smooth muscle preparations such as vas deferens, gut and arteries, and other isolated muscle preparations including coronary and skeletal. In addition it can be used for isometric tension measurements in muscle. The Multistim is a flexible modular stimulator system with over 20 components to choose from in order to configure the optimum system for each experiment. Two



Multistim — the multichannel tissue bath stimulation system.

case sizes — a 19-inch rack-mountable accommodating up to 10 stimulation channels and a 'half-width' benchtop case accommodating up to four channels — are available. Several stimulator channel modules can be obtained with voltage protocols up to 100 V and current capacities up to 2.5 A per channel. Overload cut-out protection is standard. Various timing and triggering modules are offered, which allow convenient implementation of pulse train and other stimulus delivery protocols.

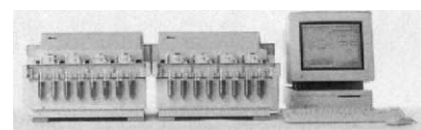
A series of biocompatible 'micro' flow-through electrodes, which are suitable for whole organ or tissue slice perfusion studies, has been introduced by Lazar Research Laboratories (*Reader Service No. 108*). These electrodes include micro flow-through modules for pH, pO_2 , pCO_2 and temperature. The probes are designed to offer researchers on a limited budget a



Biocompatible micro flow-through sensors for organ perfusion studies.

modular, cost-effective approach to real-time monitoring of organ perfusion experiments. The four microsensors are also available as a complete low-cost monitoring system (Model PFS-2). The system offers a console to organize all sensing modules, a flow-through cell tray to simplify plumbing of the flow-through electrodes, a digital real-time monitor and channel selector for independent monitoring of all parameters, and an output connector for direct hook-up to a strip-chart recorder. Interface to an IBM PC or compatible computer is also available. All sensors and system components are biocompatible and are resistant to the build-up of protein or clogging. Sensor and flow-cell components can be chemically sterilized.

The Cytosensor microphysiometer system from Molecular Devices is designed to provide functional responses of living cells in minutes and eliminate the need to develop and perform a variety of assay types in the laboratory (*Reader Service No. 109*). The company stresses that because the system detects a physiological response that is specific, understanding the signal transduction pathways is not essential in order to obtain response curves for effector/inhibitor agents.



Drug characterization and signal transduction evaluation is possible using a single assay tool with the Cytosensor microphysiometer system.

In addition, unknown signal transduction pathways can be elucidated by using various signal transduction probes to block or inhibit specific pathways. Utilizing specifically designed chambers for the containment of cultured cells, various ligand-receptor interactions can be evaluated. Effector agents can include growth factors, cytokines, signal transduction probes, muscarinic, β -adrenergic, and dopaminergic drugs, as well as many other drugs involving different receptor categories.

■ Reagents

Troponin, a contraction-regulating protein, is central to research into the mechanism and consequent diagnosis of acute myocardial infarction (*Reader Service No. 110*). Biogenesis is now offering various antibodies and antigens related to troponin which are suitable for most aspects of immunoassay-based research applications. The range includes a matched pair of monoclonals for troponin T (cardiac). Other potential uses are in immunohistochemistry and immunoaffinity purification.

Wortmannin, a fungal metabolite available from Kamiya Biomedical, is a specific and potent inhibitor of myosin light-chain kinase and a potent inhibitor of neutrophil activation (*Reader Service No. 111*). In neutrophils, wortmannin inhibits fMet-Leu-Phe (fMLP)-stimulated superoxide anion production without affecting the mobilization of intracellular calcium and inhibits fMLP-stimulated phospholipase-D activation. Kamiya Biomedical also says that wortmannin has been found to inhibit phosphatidylinositol 3-kinase and block IgE-mediated histamine release in rat basophilic leukaemia cells and human basophils. Wortmannin has no effect on cAMP-dependent protein kinase, cGMP-dependent protein kinase and calmodulin-dependent protein kinase II. Stated applications include biochemistry, cancer research, cell biology, immunology and physiology.

TCS Biologicals is distributing a new range of highly purified and active preparations of phosphatases and kinases for signal transduction research on behalf of Upstate Biotechnology (*Reader Service No. 112*). Antibodies to human cyclins, cdc2 kinase, protein phosphatase-2A purified enzyme, SH-protein tyrosine phosphatase-1 purified enzyme and casein kinase-1 purified enzyme (ELISA-based) are just a few of the new products on offer.

A selection of new **neuroligands for biomedical research** are available from Amersham (Reader Service No. 113). The line-up includes [125 I] ω -conotoxin MVIIA for brain research, [prolyl $^{2-4}$ -3,4(n)- 3 H][Pro 9] substance P, said to be the most selective NK ligand available, and [methoxy- 3 H]Glyburide, an ATP-sensitive potassium channel blocker.

Becton Dickinson Labware/Collaborative Biomedical Products is expanding its line of **proteoglycans** to include chondroitin sulphate proteoglycan (CSPG) and dermatan sulphate proteoglycan (DSPG) (Reader Service No. 114). These large, multifunctional



Becton Dickinson's new proteoglycans — chondroitin sulphate proteoglycan and dermatan sulphate proteoglycan.

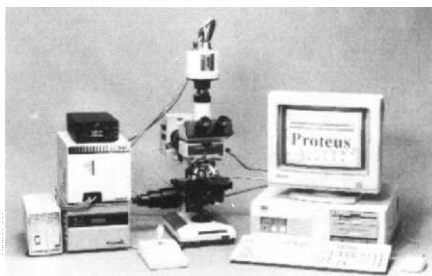
molecules with numerous unbranched glycosaminoglycan chains linked to a core protein are important components of the extracellular matrix (ECM) and are widely distributed in a variety of connective tissues. Purified proteoglycans are useful to study their interaction with cell surface receptors (for example, integrins), growth factors (fibroblast growth factor, epidermal growth factor and transforming growth factor- β), and other ECM components. Such studies will help determine the role of proteoglycans in signal transduction and ECM-mediated regulation of cell proliferation, differentiation, adhesion and migration. In addition, these products can be used as markers to define structural alterations occurring in diseases such as arthritis and atherosclerosis. The chondroitin sulphate proteoglycan and dermatan sulphate proteoglycan are isolated from bovine aorta. Both are assayed for their ability to inhibit spreading of human fibrosarcoma cells (HT-1080 cells) on plastic surfaces treated with fibronectin. CSPG and DSPG are supplied in quantities of 1 mg and 0.5 mg per vial, respectively.

Human tenascin, a hexameric, multi-domain **extracellular matrix glycoprotein** with disulphide-linked subunits of 190 to 230 kD, is now available through Gibco BRL/Life Technologies (Reader Service No. 115). It is known to be involved in the development of the central nervous system and mesenchymal-epithelial derived organs; binding of chondroitin sulphate proteoglycan, fibronectin and integrins; cell adhesion; growth and migration of tumour cells, neuronal cells and epithelial cells; and

wound healing. Gibco BRL's human tenascin has been isolated and purified from the human glioma cell line U-251MG. Based on ELISA testing, this material is said to contain less than 1 per cent laminin and no fibronectin. By SDS-PAGE analysis, the protein migrates around 280 to 300 kD. It has been tested on human cell lines and is reported to be functional on chicken and rodent cell lines. Units of 100 μ g cost US\$210.

■ Confocal configurations

Vaytek, the company that markets a **digital confocal microscope** using true PSF deconvolution algorithms, is introducing a new high-end system that includes a cooled CCD camera, computer, microscope shutter and stage stepper, framegrabber, and software for data acquisition, deconvolution and two- and three-dimensional image processing (Reader Service No. 116). Proteus features three algorithms for removing out-of-focus haze in microscope images: no neighbour, nearest neighbour and constrained iterative. In addition, the system includes software



Proteus — Vaytek's customized digital confocal microscope.

and hardware for precise control of microscope accessories for the capture of optical sections. Proteus is intended for *in situ* hybridization studies but is designed to be flexible enough to work with any light wavelength and any light mode. Vaytek says that it can be customized to work with most existing microscope and image acquisition systems.

New off the production line at Bio-Rad is the MRC 1000 **confocal imaging system** — a computer-controlled, multichannel laser scanning imaging system designed to combine an advanced applications capability with ease of use (Reader Service No. 117). Up to three independent confocal detection channels, each with computer-controlled gain and black level, are offered. The confocal iris and six-position emission filter wheel on each channel are also independently controlled by the computer. The system offers simultaneous acquisition of multiple channel data as standard; signal mixing electronics (additions and subtractions) allow combined display of multichannel data. The subtractive mixing capability provides the facility for real-time bleed through corrections. The instrument utilizes a fibre-

ADVERTISEMENTS

Highest Quality Oligonucleotides

Synthesised, purified and shipped in 3 days

£2.50 per base

(All inclusive)

0.2 μ M scale (approx. 500 μ g)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit

King's College School of Medicine & Dentistry

Tel. 071 - 326 3126 FAX 071 - 733 3877

Reader Service No.56

epitope
custom
peptides

Unit 308
Newcastle Technopole
Trafalgar Street
Newcastle upon Tyne
NE1 2LA. UK

Tel. UK 091 201 2179 Fax. 091 201 2191

- Custom peptide synthesis
- Protein Sequencing
- Antibody Production
- Computerised Epitope Prediction

FREE Polo-shirt with orders over £400

dti SMART Award Winner 1993

Reader Service No.26

Affordable Fluorescence Imaging & Photon Counting Ion Quantitation Systems

Records intracellular ion concentrations of Ca^{2+} , Na^{+} , K^{+} , Mg^{2+} , Cl^{-} , pH, etc.

- * Single Cells; * Cell Suspensions;
- * Adherent Cells; * Tissue Surface.

- * Sampling rates as high as 60 ratio images/sec
- * Dual/Single wavelength Chopper based imaging with automatic synchronization of the wavelength alternation to the video frame rate
- * Simultaneous recording of other data (cell length...)
- * User friendly software for DOS / Windows environments

ION OPTIX
CORPORATION

44 Parkway Crescent
Milton, MA 02186 USA
Ph/fax: (617) 696-7335

Reader Service No.13

NEUROSCIENCE ANTIBODIES

- ✓ Choline Acetyltransferase
- ✓ Dopamine Receptor D2/D3
- ✓ Glutamate Receptors
- ✓ Neurofibrillary Tangles
- ✓ Oligodendrocytes
- ✓ Tau
- ✓ Tyrosine Hydroxylase

For a complete listing of our products please call, fax or write us today.

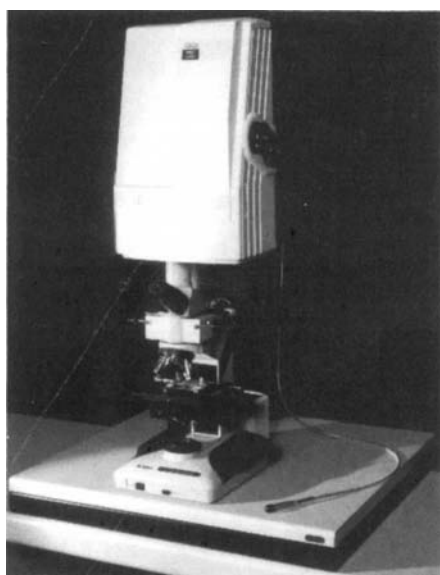
800-437-7500



CHEMICON International, Inc.

28835 Single Oak Drive, Temecula, CA 92590 USA
Phone 909-676-8080 • Fax 909-676-9209

Reader Service No.12



Bio-Rad's MRC 1000 computer-controlled confocal imaging system.

optic delivery system with computer-controlled selection of laser intensity and wavelength, without change of beam pointing accuracy. A motorized pillar assembly that provides quick changeover from an upright to an inverted configuration is an optional extra.

■ Assorted antibodies

Biocytex offers two **monoclonal antibodies that recognize neuron structures** on human, pig, rat and mouse tissues (*Reader Service No. 118*). One of the monoclonal antibodies recognizes cell bodies and axons; the second recognizes only nerve processes. On frozen sections of nerve tissues, these two antibodies strongly stain neuronal structures from the cortex to the neuromuscular junction.

Chemicon offers a wide variety of antibodies and antigens for use in **integrin and extracellular matrix protein research** — details of which are contained in the company's new brochure, "*Tools for Integrin and ECM Research*" (*Reader Service No. 119*). Among those highlighted are polyclonal antibodies to integrins β 1 and β 4 and fibronectin receptor. Monoclonal antibodies include those to integrins α IIb, α 6, α V, β 1 and β 3. Also featured are monoclonals to collagen type VI, laminin B1 chain and B2 chain and entactin/nidogen.

Comparative studies of mesothelioma markers using Leu-M1, CEA and thrombomodulin have demonstrated thrombomodulin to be an effective tool for research into mesotheliomas. Studies by Fink *et al.* (*Int. J. devl Biol.* 37, 221; 1993) have shown thrombomodulin to be present on most mesotheliomas but absent from most adenocarcinomas of the lung. Therefore, Dako's new **anti-thrombomodulin monoclonal**

antibody may be useful for the research of such tumours (*Reader Service No. 120*). It may also be useful for researching the role that thrombomodulin plays in the repair process following a vascular injury because of the relationship between thrombomodulin and protein C. (Protein C, which functions as an anticoagulant when combined with protein S, is activated by its interaction with thrombin and thrombomodulin and contributes to the prevention of thrombosis on the surfaces of endothelial cells. A 75-kD transmembrane glycoprotein, thrombomodulin has a normal distribution that includes the lining of blood and lymphatic vessels, mesothelial cells and some macrophages of the lung, meningeal lining cells syncytiotrophoblasts, megakaryocytes and platelets. It has also been found in a subset of islet cells and in peripheral nerves.

Genzyme's **polyclonal antibody to human interleukin-1 receptor antagonist** should be of interest to immunologists and researchers in inflammatory disease (*Reader Service No. 121*). Raised in rabbits, this antibody recognizes both the intracellular and extracellular forms of human IL-1 receptor antagonist. It can be used in radioimmunoassays, western blots, immunocytochemistry and neutralization techniques, in addition to its use as either a capture reagent or second antibody in ELISAs.

■ Microdialysis analyses

New from CMA/Microdialysis — a **microdialysis pump and fraction collector** (*Reader Service No. 122*). The CMA/102 pump is a compact dual-syringe device that is precalibrated for 1 ml or 2.5 ml syringes. Each syringe is controlled separately, allowing the pump to be used as two independent units. For more advanced protocols, the pump can be connected to a computer through an



CMA/Microdialysis's new CMA/142 fraction collector collects samples from 1 μ l to 50 μ l.

RS232 interface. The CMA/142 microfraction collector can be used as a stand-alone unit to collect samples from 1 μ l to 50 μ l from one or two microdialysis probes.

ESA announces the availability of a new application note on **acetylcholine in rat brain microdialysis perfusates** (*Reader Service No. 123*). This application, which uses ESA's optimized acetylcholine analysis system, is said to provide detection limits of less than 30 fmols. With this microdialysis method, it is possible to dialyse discrete brain regions with minimal disruption of vital functions by minimizing the need for high levels of acetylcholine esterase inhibitors. The note describes the method, results and optimized parameters necessary to see basal acetylcholine levels in rat brain. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

**Teflon*
Coated
and
Bare Wire from**



MEDWIRE

Medwire provides the highest quality platinum, platinum alloy, silver and stainless steel wire in both single and multistrand configurations.

For prompt delivery or additional information, please contact:

MEDWIRE, an affiliate of Sigmund Cohn Corp.
121 So. Columbus Ave. Mount Vernon, NY 10553
Tel: (914) 664-5300 Fax: (914) 664-5377

*Reg. TM of E. I. DuPont & Co

Reader Service No.25

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG